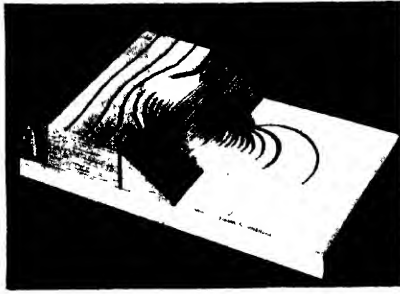


Prof- A K. SAHA
NUCLEAR PHYSICS DIVISION.
SAHA INSTITUTE OF NUCLEAR PHYSICS.
92, Acharya Pratulla Chandra Road,
CALCUTTA-9.

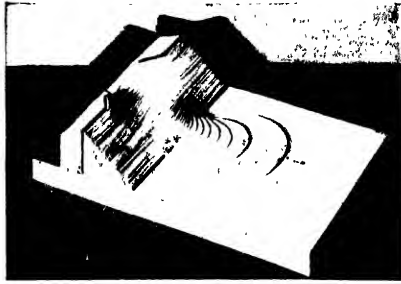
DR. AJIT KUMAR SAHA
125, SOUTHERN AVENUE,
CALCUTTA-29

**INTRODUCTION TO
ELECTRON OPTICS**

**Prof:- A K. SAHA.
NUCLEAR PHYSICS DIVISION.
SAHA INSTITUTE OF NUCLEAR PHYSICS.
82, Acharya Prafulla Chandra Road,
CALCUTTA-9.**



a. Immersion lens (grid positive). (*Bruche and Scherzer.*)



b. Immersion lens (grid negative). (*Bruche and Scherzer.*)



c. Symmetrical lens, with central element negative.
(*Marconi's Wireless Telegraph Co. Ltd.*)

PLATE I. GRAVITATIONAL MODELS OF POTENTIAL FIELDS.

INTRODUCTION
TO
ELECTRON OPTICS

THE PRODUCTION
PROPAGATION AND FOCUSING OF
ELECTRON BEAMS

BY

V. E. COSSLETT

M.A. (OXON.), M.SC. (LOND.), PH.D. (BRISTOL)

Prof. A. K. SAHA.

NUCLEAR PHYSICS DIVISION.

SAHA INSTITUTE OF NUCLEAR PHYSICS,

92, Acharya Prafulla Chandra Road,

CALCUTTA-9.

OXFORD
AT THE CLARENDON PRESS

1946

OXFORD UNIVERSITY PRESS
AMEN HOUSE, E.C. 4
LONDON EDINBURGH GLASGOW NEW YORK
TORONTO MELBOURNE CAPE TOWN BOMBAY
CALCUTTA MADRAS
GEOFFREY CUMBERLEGE
PUBLISHER TO THE UNIVERSITY

PRINTED IN GREAT BRITAIN AT THE UNIVERSITY PRESS, OXFORD
BY JOHN JOHNSON, PRINTER TO THE UNIVERSITY

PROF. A. L. SAHA
NUCLEAR PHYSICS DIVISION
SAHA INSTITUTE OF NUCLEAR PHYSICS,
92, Acharya Pratulla Chandra Road,

CALCUTTA-8
PREFACE

AS a consequence of the development in recent years of the cathode-ray tube, new types of thermionic valve, and many related devices, a physicist is now introduced very early in his career to the controlled use of electron beams. Such apparatus is not only widely employed in industry and the Services, it has also become a necessity for research and advanced teaching in every laboratory. As always it is desirable that the student or research worker should be aware of the principles on which his instruments are based, but in this field theoretical development has lagged behind application, many electronic devices having been conceived and practically realized long before a fundamental explanation of their operation was available. A position has now been reached, however, where an adequate explanation of most of the phenomena concerned may be given in familiar terms and the path to a more comprehensive treatment indicated.

It has thus become possible, and important for the training of the applied physicist in particular, to include in an Honours Physics Course the subject of Electron Optics: the study of the production, propagation, and focusing of electron and ion beams. This book is the result of a series of lectures given at Oxford to undergraduates in their final year. The course has now been repeated four times and has accommodated itself progressively to its audience. It was early obvious that none of the existing works on Electron Optics was suitable for the average student. The short monograph by Klemperer is excellent for physicists in the research stage, who already know their way in their subject, but it is too concise for beginners. On the other hand, Myers's book is too compendious for them, although very valuable for reference. Picht's treatment, apart from the language difficulty, is highly theoretical and omits all discussion of practical applications; it is commendable primarily for its mathematical rigour. The present text has therefore been prepared as an introduction to the subject, intermediate in length and level of treatment, in the hope that it may meet satisfactorily a growing need in universities and research laboratories. It has drawn considerably upon the work of the authors mentioned, to whom the writer desires here to express his acknowledgement.

Throughout the book an effort has been made to keep mathematical exposition subordinate to the description of physical principles. Nevertheless the investigation of electron lenses has now so far progressed that a full treatment of aberrations cannot be omitted, and this necessarily involves more complicated mathematical discussion. A good grounding in the calculus is accordingly required of the earnest student, including some acquaintance with partial differential equations and preferably a knowledge of the nature of Bessel's functions. At a first reading, however, a sufficient survey of the subject may be obtained if sections 7, 12 (d) and 36, and all but the first section of Chapter V be omitted; the relevant sections

are marked with an asterisk (*). Duplication of material has not been avoided where it seemed likely to smooth the path of the indifferent mathematician, especially when an approximate treatment could be given before the full theoretical discussion, bringing out more clearly essential physical principles.

A comprehensive treatment of Electron Optics, and the most elegant approach to aberrations of the image, is afforded by Hamiltonian methods; but a knowledge of these can hardly be demanded of those to whom this 'Introduction' is addressed, and the author has been content to indicate the main principles in an Appendix. The body of the text employs the trigonometrical approach, which is closely related to the usual methods of light optics.

Opportunity has been taken to widen the scope of the lecture course by adding descriptions of the more important electronic apparatus as well as by more detailed discussion of theory. In this laboratory a separate course of lectures on the cathode-ray tube follows that on Electron Optics, and moreover there now exist a number of excellent introductory manuals on this instrument. Hence only the fundamentals of its construction and operation are dealt with here, together with a short sketch of its origin and an account of some special types of tube.

The literature of Electron Optics is large and rapidly growing. In order to assist the student in more detailed exploration than this book can afford, a list of 'Further Reading' has been given at the end of each chapter, in addition to the Bibliography of major works. In selecting references for inclusion in these lists the criterion has been conciseness and clarity of treatment rather than priority of publication; whenever available a summarizing article has been quoted, such as appear in the annual *Reports on Progress in Physics*. No attempt has been made to exclude references in German or other foreign languages, but for the convenience of the average student a source in English has been included wherever possible.

It is at the same time a pleasure and an obligation to express my thanks to all those at the Electrical Laboratory who have in greater or less degree assisted with the production of this work: in particular to R. T. Lattey and M. S. Wills, who have read parts of the manuscript, for constructive criticism by which a number of errors have been corrected and obscurities removed; to Dr. F. B. Pidduck and Dr. A. v. Engel, for helpful discussion on particular points; to Dr. R. Sas-Kulczycki and Dr. H. Motz (of the Engineering Laboratory, Oxford), for permission to include unpublished work; and to my wife, Dr. A. Cosslett, for continuous aid and encouragement, without which the book could hardly have been written.

In a work of this type, prepared amid the preoccupations of war-time, errors are only too probable of occurrence. The author will be greatly obliged for information of such as come to the notice of discerning readers.

V. E. C.

ELECTRICAL LABORATORY, OXFORD

April 1945

ACKNOWLEDGEMENTS

THE author takes this opportunity of expressing his thanks to various authors, societies, and publishers for permission to use diagrams and plates which have appeared in the following books and periodicals:

Applied Electronics (John Wiley and Sons, Inc.) for Fig. 87.

C. H. Bachman and S. Ramo and *Journal of Applied Physics* for Figs. 127 *b* and 128.

E. Brüche and O. Scherzer, *Geometrische Elektronenoptik* (J. Springer, Berlin), for Plates I *a* and *b*.

Messrs. A. C. Cossor, Ltd., London, for Plate V *a*.

Prof. G. I. Finch and *Reports on Progress in Physics* (London), for Figs. 118 and 119.

Prof. G. I. Finch, A. G. Quarrell, and H. Wilman, for Plates IV *a*, *b*, *c*, and *d*.

Dr. D. Gabor and *Electronic Engineering* for Figs. 78 and 80.

L. S. Goddard and O. Klemperer, and *Proceedings of the Physical Society of London* for Fig. 61.

Dr. P. Grivet and Cie Gle de T.S.F., Paris, for Plates V *b*, and VI *b* and *c*.

M. E. Haine and Metropolitan Vickers Ltd., Manchester, for Fig. 123 and Plate VIII *c*.

Dr. R. D. Heidenreich for Plate VII *a*.

Dr. R. A. Houston, *Introduction to Mathematical Physics* (Longmans, Green and Co., Ltd.), for Fig. 11.

Dr. H. Iams and Dr. H. Rose, and *Proceedings of the Institute of Radio Engineers* (of America), for Fig. 108.

Dr. R. O. Jenkins and Messrs. G. E. C. Ltd., London, for Plates III *a*, *b*, *c*, and *d*.

Dr. O. Klemperer, *Electron Optics* (Cambridge University Press), for Figs. 18, 33, 35, 42, 46, 47, 66, 67, 143, 145, 146 *a*.

Dr. O. Klemperer and H. Miller, and *Journal of Scientific Instruments*, for Fig. 55.

Dr. O. Klemperer and W. D. Wright, and *Proceedings of the Physical Society of London*, for Fig. 32.

Dr. J. D. McGee and Dr. H. G. Lubzynski, and *Journal of the Institution of Electrical Engineers*, for Fig. 107.

Messrs. Mullard Radio Valve Co. Ltd., London, for Plates II *a* and *b*.

L. M. Myers, *Electron Optics* (Chapman and Hall, Ltd.) for Figs. 15, 23, 34, 36, 40, 48, 52, 112 *a*, 144 and Plates I *a*, *b*, and *c*.

G. Parr, *Cathode Ray Oscillograph* (Chapman and Hall, Ltd.), for Fig. 112 *b*.

E. Pollard and W. L. Davidson, *Applied Nuclear Physics* (John Wiley and Sons, Inc.), for Fig. 142 *a*.

Radio Corporation of America for Plate VI *a*;

J. Smiles, and the National Institute of Medical Research, London, for Plates VII *b* and VIII *d*.

F. E. Terman, *Fundamentals of Radio* (McGraw-Hill Book Company, Inc.) for Fig. 146 *b*.

Ultra-High Frequency Techniques (Chapman and Hall, Ltd.), for Figs. 113, 136, 138, 150, 151, 152, 153, 154.

Dr. V. K. Zworykin and J. Hillier, and *Journal of Applied Physics*, for Fig. 124.

Dr. V. K. Zworykin, J. Hillier and R. L. Snyder, and *Journal of Scientific Instruments*, for Fig. 130.

Dr. V. K. Zworykin and R. Morton, *Television* (John Wiley and Sons, Inc.) for Fig. 39.

Dr. V. K. Zworykin and J. A. Rajchman, and *Proceedings of the Institute of Radio Engineers* (of America), for Figs. 100 and 101.

The electron micrographs in Plates VIII *a* and VIII *b* were taken at the National Physical Laboratory, and are reproduced by courtesy of the Director and by permission of the Controller of His Majesty's Stationery Office. Crown Copyright is reserved.

CONTENTS

List of Plates	xi
Abbreviated Titles of Journals to which Reference is made	xii
I. INTRODUCTORY	
1. General	1
2. Early Experimental Work	3
3. Deflexion in a Transverse Magnetic Field	4
4. Deflexion in a Transverse Electric Field	5
5. Deflexion in Coincident Electric and Magnetic Fields	6
6. Fields Crossed at Right Angles; Determination of e/m	7
*7. Crossed Fields: Electrons emitted from Plate	7
8. Refraction of Electron Beams	9
9. Limitations of the Optical Analogy	11
10. Beam Parameters	13
II. THE ELECTROSTATIC FIELD	
11. General	15
12. Determination of the Potential Field	16
13. Direct Calculation of Field Distribution by the Relaxation Method	22
14. Experimental Determination of the Potential Field	25
15. Motion of Electrons in a Potential Field	29
16. Trigonometrical Ray-tracing	30
17. Ray-tracing by Other Methods; Automatic Tracing	32
18. Deduction of Trajectories from Axial Potential Distribution	34
19. Direct Determination of Trajectories with a Gravitational Model	38
III. ELECTROSTATIC FOCUSING	
20. Types of Lenses	42
21. Cardinal Points	44
22. Cardinal Points of Electron Lenses	48
23. Experimental Location of Cardinal Points	49
24. The Two-cylinder Lens	52
25. Aperture Lenses	56
26. Symmetrical and Univoltage Lenses	62
27. Immersion Lenses	68
28. Electron Mirrors	71
IV. MAGNETIC FOCUSING	
29. Introductory	77
30. Solenoidal Lenses	77
31. The Short Magnetic Lens	80
32. Form of Field in Short Magnetic Lens	87
33. The Magnetic Immersion Objective	91
34. Ray-tracing in Magnetic Lenses	97
35. Magnetic Vector Potential; Equivalent Electrostatic Lens	101
*36. Superimposed Magnetic and Electrostatic Fields: General Case	105

* May be omitted from a first reading.

CONTENTS

V. ABERRATIONS OF THE IMAGE

37. Introduction	113
*38. General Treatment of Aberration	117
*39. Spherical Aberration	121
*40. Distortion	130
*41. Curvature of the Field and Astigmatism	131
*42. Coma	136
*43. Chromatic Aberration	138
*44. Correction by Space Charge; the Ideal Lens	142

VI. THE PRODUCTION OF ELECTRON BEAMS

45. Electrons in Metals	148
46. Thermionic Emission	153
47. The Electron Gun	155
48. Field Emission	160
49. Field Emission Microscope	163
50. Photo-electric Emission	164
51. Image Transformer Tubes	166
52. Secondary Emission	168
53. Electron Multipliers	169

VII. THE CATHODE-RAY TUBE AND ITS DERIVATIVES

54. The Cathode-ray Tube	173
55. Television Transmitting Tubes	177
56. The Deflexion System	184
57. High-speed Oscillographs	190

VIII. ELECTRON DIFFRACTION: THE ELECTRON MICROSCOPE

58. Electron Diffraction: Principles	194
59. Electron Diffraction: Practice	197
60. Electron Microscopy	199
61. The Magnetic Electron Microscope	206
62. The Electrostatic Electron Microscope	210
63. Other Types of Microscope	212

IX. OTHER APPLICATIONS: CYLINDRICAL FIELDS

64. Beta-ray Spectrometry	215
65. The Magnetron	219
66. The Cyclotron	222
67. The Betatron, or Induction Accelerator	226
68. Radio Valves	229
69. Mass Spectrographs	233

X. VELOCITY-MODULATED BEAMS

70. Principles	238
71. The Klystron	242
72. Other Velocity-Modulated Tubes	246

Appendix: The Hamiltonian Method in Electron Optics	249
---	-----

General Bibliography	263
--------------------------------	-----

Tables of Electron Constants	264
--	-----

List of Symbols	266
---------------------------	-----

Index	270
-----------------	-----

* May be omitted from a first reading.

LIST OF PLATES

- PLATE I. GRAVITATIONAL MODELS OF POTENTIAL FIELDS . . . *frontispiece*
 a. Immersion lens (positive grid).
 b. Immersion lens (negative grid).
 c. Symmetrical lens, with central element negative.
- PLATE II. RUBBER MODEL OF POTENTIAL FIELD IN A (PLANAR) PENTODE
facing page 40
 a. The ball-release mechanism (= cathode), and the three rows of grid 'wires'.
 b. General view, showing arrangements for photography of tracks. The 'anode' bar is also visible.
- PLATE III. POINT-PROJECTOR, OR EMISSION, MICROSCOPE . . . *facing page 164*
 a. Image of etched molybdenum point cathode.
 b. Molybdenum cathode after heating to 1520° K. with field applied.
 c. Tungsten cathode fully activated with thorium.
 d. Thorium on molybdenum cathode, after heating to 1700° K.
- PLATE IV. ELECTRON DIFFRACTION PATTERNS . . . *facing page 196*
 a. Cross-grating pattern of hexagonal structure (mica).
 b. Cross-grating pattern of cubic structure (aluminium).
 c. Ring pattern from chromium deposited on iron, showing partial orientation (strongly pronounced (111) rings).
 d. Ring pattern from chromium deposited on bismuth, showing random disposition of crystals (very imperfect (111) orientation).
- PLATE V a. Plate structure of double-beam cathode ray tube. }
 b. Electrostatic electron microscope. } *facing page 206*
- PLATE VI a. Console model magnetic electron microscope. }
 b. Magnesium oxide smoke ($\times 8,000$). } *facing page 207*
 c. Etched aluminium surface (oxide replica) ($\times 8,000$).
- PLATE VII a. Cleavage face in fractured magnesium crystal (silica replica) ($\times 6,500$). }
 b. Malarial parasite (*Plasmodium cynomolgi*) in a red corpuscle of monkey blood ($\times 15,000$). } *facing page 208*
- PLATE VIII a. Aucuba mosaic virus ($\times 14,000$). }
 b. Diatom: *Amphipleura Pellucida* ($\times 10,000$). } *facing page 209*
 c. Zinc oxide smoke ($\times 20,000$)
 d. Bacteria (*Rickettsia*) of epidemic typhus (Cairo strain) ($\times 15,000$).

ABBREVIATED TITLES OF JOURNALS TO WHICH REFERENCE IS MADE

<i>Amer. J. Phys.</i>	American Journal of Physics.
<i>Ann. Phys.</i>	Annalen der Physik.
<i>Ark. Mat. Astr. Fys.</i>	Arkiv för Matematik, Astronomi och Fysik.
<i>Bell. Syst. Techn. J.</i>	Bell System Technical Journal.
<i>Bull. Amer. Soc. Test. Mater.</i>	Bulletin of the American Society for Testing Materials.
<i>Canad. J. Res.</i>	Canadian Journal of Research.
<i>Elect. Commun.</i>	Electrical Communication.
<i>Electronic Engng.</i>	Electronic Engineering.
<i>Electronics</i>	Electronics.
<i>G.E.C. Journ.</i>	General Electric Company Journal.
<i>J. Appl. Phys.</i>	Journal of Applied Physics (formerly 'Physics').
<i>J. Franklin Inst.</i>	Journal of the Franklin Institute.
<i>J.I.E.E.</i>	Journal of the Institution of Electrical Engineers.
<i>J. Opt. Soc. Amer.</i>	Journal of the Optical Society of America.
<i>J. R. Micr. Soc.</i>	Journal of the Royal Microscopical Society.
<i>J. Röntgen Soc.</i>	Journal of the Röntgen Society.
<i>J. Sci. Inst.</i>	Journal of Scientific Instruments.
<i>Mon. Math. Phys.</i>	Monatshefte für Mathematik und Physik.
<i>Nature</i>	Nature.
<i>Philips Techn. Rev.</i>	Philips Technical Review.
<i>Phil. Mag.</i>	Philosophical Magazine.
<i>Physica</i>	Physica.
<i>Physics</i>	Physics (now J. Appl. Phys.).
<i>Phys. Rev.</i>	Physical Review.
<i>P.I.R.E.</i>	Proceedings of the Institution of Radio Engineers (U.S.A.).
<i>Proc. Camb. Phil. Soc.</i>	Proceedings of the Cambridge Philosophical Society.
<i>Proc. Phys. Soc.</i>	Proceedings of the Physical Society (London).
<i>Proc. Roy. Soc.</i>	Proceedings of the Royal Society (A).
<i>Q. J. Math.</i>	Quarterly Journal of Mathematics.
<i>Rev. Sci. Inst.</i>	Review of Scientific Instruments.
<i>Television</i>	Television (now Journal of the Television Society).
<i>Trans. Far. Soc.</i>	Transactions of the Faraday Society.
<i>Z. Phys.</i>	Zeitschrift für Physik.
<i>Zs. f. Instk.</i>	Zeitschrift für Instrumentenkunde.
<i>Z. techn. Phys.</i>	Zeitschrift für technische Physik.

CHAPTER I

INTRODUCTORY

1. *General*

ELECTRON optics is that branch of electronics which deals with the production, propagation, and focusing of beams of ions and electrons. It may be regarded as 'internal' electronics in contrast to the 'external' electronics of circuit engineering. The subject is sometimes also referred to as 'electron ballistics', but the title of electron optics now generally preferred is justified not only by the formal similarity of much of the subject to light optics but also by the de Broglie wave properties of electron beams. The deflexion and focusing of a beam is effected by electric and magnetic fields produced by suitable arrangements of electrodes and magnets, which are referred to as 'electron lenses'. Their focusing properties, as distinct from the details of their construction, are treated in essentially the same way as glass lenses in the study of light; they are found to obey the ordinary lens formulae and to suffer from similar aberrations.

The study of the path of a charged particle in an electromagnetic field began with the first experiments on cathode rays in the 1890's, reference to which will be found in the next section. Although considerable practical developments took place during the succeeding thirty years, it was not until 1926 that the general theory of the motion of ions and electrons was formulated by Busch. Prompted by de Broglie's hypothesis, made two years earlier, he laid the basis of electron optical theory and derived the conditions in which electric and magnetic fields, separately or together, could exert a focusing effect on trajectories and thus act as electron lenses. Although his treatment is based on the wave analogy, it employs the familiar methods of classical mechanics and not wave mechanics. This book follows in the main Busch's treatment, as further elaborated by Scherzer. It uses the same line of reasoning as the usual treatment of geometrical optics in the study of light, and is consequently spoken of as the theory of geometrical electron optics.

The development of a wave theory of electron optics was due to Glaser (1932-8), who utilized Hamiltonian methods. A century earlier Hamilton had worked out the theory of the passage of light through anisotropic heterogeneous media and a closely related treatment of particle trajectories, employing the 'characteristic functions' now named after him. The electromagnetic field acts as a medium of this type towards a charged particle, since the deviating effect or 'refractive index' is found to depend on the direction as well as the position of the particle. Glaser was therefore able to transfer Hamilton's treatment directly to the electron optical case. A concise and elegant treatment of lens aberrations is then possible (see

Appendix), in contrast to the laborious method of Scherzer given in Chapter V.

Although the electron mechanical treatment followed in the body of the book was inspired by de Broglie's hypothesis, very little use need be made of his theory in the discussion. It is helpful, however, to have in mind the general nature of the analogy between particle and wave motion, which derives from Hamilton's idea of the equivalence of the passage of a light ray through an optical medium and that of a mass point through a potential field. Otherwise, except in regard to electron diffraction, the wave properties of the electron are invoked only in the discussion of certain phenomena which can so far be satisfactorily treated from the corpuscular standpoint.

The growth of electron optical theory has been closely bound up with the development of its applications. The empirical construction of electron lenses stimulated theoretical investigation, and on the other hand deductions from theory led to the rapid improvement of focusing systems. Electrostatic and magnetic lenses were used to improve greatly the early cathode-ray tube, to construct electron microscopes of higher resolution than the best light microscopes, and to develop television apparatus for transmitting and reproducing images of high definition. The diffraction of electrons gave valuable information on the structure of crystals and of molecules in the gaseous state, where X-rays could not be employed. The design and performance of thermionic valves and X-ray tubes have been improved, and special types of high-frequency oscillators developed. Other applications have been made in atomic physics, especially to beta-ray spectrometry and to the design of high-voltage machines such as the cyclotron and betatron. In a very short time, therefore, electron optics has arisen as a new and extensive branch of physics; it is barely twenty years old, and there can be little doubt that the controlled use of electron beams is only in its beginnings.

The major applications mentioned above and a number of others are described in later chapters, the first half of the book being devoted to theoretical exposition. The source of the electron beam is an important part of any application and usually involves special electron optical problems of its own. A chapter has therefore been devoted to the discussion of the various means of producing electrons; as in light, the production, propagation, and focusing of the beam all form part of the study of the subject. In many places the optical analogy will be directly invoked, and throughout the terms 'ray', 'path', and 'trajectory' are used synonymously. In the same way, in the discussion of deflexion and focusing, the word 'electron' must be understood to include 'ion', with the appropriate change in sign of the charge and of the direction of the electric and magnetic force on a positive ion; in a number of places special problems concerning the use of ion beams have also been treated. In the general discussion of theory a particle of positive charge has been assumed, but in many special cases, especially of applications where electrons alone

come in question, the negative sign for the charge (e) is inserted; this should be clear from the context even when no specific statement to the effect is made.

2. *Early Experimental Work*

In the fourth quarter of last century great interest was taken in the phenomena of electrical discharges between two metal electrodes maintained at a high potential difference in a partially evacuated tube. Observation showed that, under conditions of very low pressure, a well-defined beam was produced, which was propagated apparently in a rectilinear manner and which transmitted momentum to solid objects in its path. Crookes, Perrin, and J. J. Thomson showed that these 'cathode rays' were essentially the same whatever metal was used in the electrodes, that

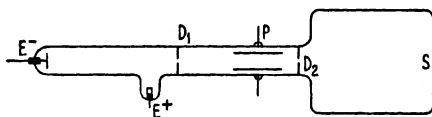


FIG. 1. Braun cathode-ray tube.

they carried a negative charge and could be deflected by both magnetic and electric fields. The term 'electron' was first used for the individual corpuscles by Johnstone Stoney in 1890, and the first practical 'cathode-ray tube' for their investigation was devised by Braun in 1897 (Fig. 1). This was the prototype of the apparatus used by Thomson, Townsend, and others in subsequent experiments.

It consists of two electrodes ($E+$, $E-$), usually of aluminium, placed some 20 cm. apart in a glass envelope carrying at its farther end a coating of fluorescent material such as calcium tungstate or zinc sulphide. The tube is evacuated and sealed off permanently, the residual pressure being of the order of 10^{-3} mm. Hg only, and a potential difference of some tens of thousands of volts is applied between the electrodes. Under this intense field (see Chapter VI) electrons leave the cathode and pass down the tube towards the anode. The field strength being very high near the surface of the disk cathode, the electrons are accelerated rapidly to a very high velocity in practically straight-line paths and are very little affected by the curved field near the off-set anode. They are limited to a beam of small cross-section by mica diaphragms D_1 , D_2 , and pass between a pair of metal plates P before striking the screen S , where they give rise to a bright spot of light. An electric field applied between the plates P will result in the deflexion of the (negatively charged) beam, and hence of the light spot, in the direction of the positively charged plate. A transverse magnetic field applied in the same region gives rise to a deflexion at right angles to the axes of both beam and magnetic field, as required by elementary electro-magnetic theory ('Left-hand Rule'). By applying magnetic and electric fields in opposing directions and of such magnitude that

the resulting deflexion of the beam was zero, J. J. Thomson was able to determine the ratio of charge to mass for the electron (cf. p. 7).

This type of tube was used for many years in experimental work, being improved by the substitution of a heated oxide-coated cathode (1905) and an electrostatic concentrating electrode around the cathode (1908), both due to Wehnelt. Wiechert in 1899 observed the concentrating effect on the beam-spot of a short axial magnetic field produced by a current-carrying coil wound round the tube, and this form of focusing was subsequently much used, but the reasons were neither appreciated nor investigated. Otherwise the tube remained substantially unchanged until Zworykin's work in 1928-9 (cf. Chapter VII).

3. Deflexion in a Transverse Magnetic Field

The passage of an electron beam through a transverse magnetic or electric field results in a deflexion which can readily be treated. In the first case, suppose electrons moving with velocity v to enter a trans-

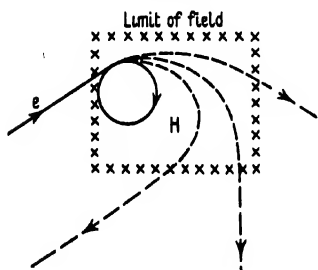


FIG. 2. Circular paths of electrons in a transverse magnetic field H (directed into the diagram).

verse magnetic field of strength H as in Fig. 2, where the field lines are assumed to be perpendicular to the plane of the diagram. (Here, as elsewhere in this book, we assume a perfect vacuum to exist, so that $H = B$, the magnetic induction.) If the electronic charge is e , the beam will be equivalent to a current evn , when vn electrons pass per second, and the magnetic force on the current will be $Hevn$. On a single electron the force is Hev , at right angles to both field and the original direction

of motion of the electron, which will thus be deflected in the plane of the diagram and continue to intersect the field lines at right angles. The electron in this way experiences a constant force in a direction perpendicular to its motion at any moment, and therefore describes in the field a circular path of radius r determined by the equivalence of the centripetal and deflecting forces:

$$Hev = mv^2/r,$$

where m is the mass of the electron, or

$$Hr = mv/e, \quad (I.1)$$

the so-called electronic momentum. This quantity, being readily determinable by direct experiment, is frequently used to specify the velocity of an electron beam.

If the field is uniform and sufficiently great in relation to the velocity of the electron, the resulting path will be a closed circle in which the electron will continue to revolve until disturbed by some other force.

This type of circular deflexion is utilized in the operation of the magnetron valve, the cyclotron, and other devices (see Chapter IX). If the field is limited in extent, the electronic path will be a more or less sharply curved arc of a circle, according to the strength of the field; a number of possible trajectories are shown in Fig. 2. If the field is not uniform, the path will be of continuously varying radius and can then only be determined by a point-to-point evaluation, based on detailed exploration of the field.

Again, if the field is uniform and the incident electron does not enter it normally but with a component of velocity in the direction of the field, the path will be more complex. The velocity component normal to the field will give rise to a rotational deflexion, and the projection of the path on a plane at right angles to the field will be a circle. In the absence of an electrostatic field the velocity along the field will be unaltered, causing the actual path to be a helix of constant pitch in the surface of a cylinder of radius given by relation I.1, if the normal velocity component is written for v (Fig. 4, full line).

4. Deflexion in a Transverse Electric Field

Suppose now the electron beam to enter the space between two charged parallel plates P_1 , P_2 (Fig. 3), in the absence of a magnetic field. If a

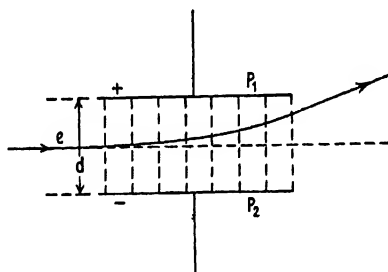


FIG. 3. Parabolic path of an electron in a transverse electrostatic field.

potential difference V is maintained between them and their normal separation is d , the electric field E in the gap will be

$$E = V/d.$$

The force on the moving electron will be given by the product Ee , acting transversely to its initial path, which is assumed to be parallel to the plates. It will therefore experience a constant acceleration in this direction,

$$a = Ee/m,$$

and will suffer a transverse displacement given by $\frac{1}{2}at^2$, if t be the time measured from the instant of entry into the field. The forward velocity of the electron will be unchanged, and its path in the field will thus be

parabolic; its motion is entirely analogous to that of a material particle in a uniform gravitational field.

The electric field between the plates will not, in fact, be uniform, mainly on account of distortion at the edges of the plates ('fringing'). The deflection in, for instance, a cathode-ray tube consequently departs slightly from the theoretically parabolic path (cf. section 56).

5. Deflexion in Coincident Electric and Magnetic Fields

When an electron is exposed to the simultaneous action of electric and magnetic fields the total force on it will be the sum of the forces exerted by each field independently. In general they will not be in the same direction and the addition must be made vectorially, remembering that the magnetic force is perpendicular both to the initial path of the electron and the magnetic field. The total force F on the electron may then be expressed as

$$F = e(E + v \times H) = ma,$$

e and E being in electromagnetic units. The resultant path of the electron is difficult to evaluate except when the directions of the two fields are simply related. In doing so, it must be remembered that the magnetic field does not affect the kinetic energy of the electron, since its effect is always perpendicular to the electron path. Hence any changes in velocity and kinetic energy in crossed fields are due to the electric field alone. In general, as the velocity changes the radius of the path described in the magnetic field varies accordingly (see equation I.1), and the path will not be circular.

The simplest configuration is that in which both fields have the same direction as the moving electron. It is then unaffected by the magnetic

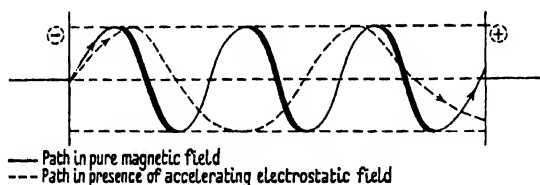


FIG. 4. Spiral paths of electrons in concurrent electric and magnetic fields.

field and suffers uniform acceleration in the electric field. If the fields are parallel but the electron has an initial velocity component perpendicular to their direction, it will move in a circular orbit under the action of the magnetic field whilst being accelerated in the electric field. The resultant path will then be a helix with a pitch that increases or decreases with time according as the acceleration is positive or negative. Since the perpendicular component of velocity is unchanged by either field, the diameter of the helix remains constant; that is, the electron describes a path on the surface of a cylinder of radius given by equation I.1 (see Fig. 4, broken line).

6. Fields Crossed at Right Angles: Determination of e/m

Referring again to Fig. 3, we now suppose both a magnetic field perpendicular to the plane of the diagram and the electric field between the plates P_1 , P_2 , to act together on an electron beam entering parallel to the plates and normal to the direction of the magnetic field. The respective deflexions will be in the same plane (that of the diagram) and may reinforce or counteract each other, according to the relative directions of the fields. If the upper plate is positively charged, and the magnetic field is directed into the diagram, their action on the electron beam will be in opposite senses. Either field may then be varied until the resultant deflexion is zero.

This null arrangement of orthogonally crossed fields was originally devised by J. J. Thomson in order to determine the ratio of the electronic charge and mass. The electric deflecting force is Ee , the magnetic force is Hev ; in the critical condition when no deflexion results,

$$eE_c = evH_c, \quad \text{or} \quad v = E_c/H_c, \quad (\text{I.2})$$

where the subscript c indicates the critical value of the fields. By accurate measurement of the fields the value of v , the velocity of the electron, may thus be found and substituted in expression I.1, giving

$$e/m = E_c/rH_c^2, \quad H_c = mv/c \cdot E/H \quad \text{or} \quad H_c = m_e \cdot E/H$$

where r , the radius of curvature of the electron path in the critical field H_c , is known from observations in the magnetic field alone. By this means Thomson found a value of 7.7×10^6 e.m.u./gm. for e/m , which is in the same order of magnitude as the result of modern work: 1.758×10^7 e.m.u./gm. From experiments which will not be described the charge on the electron is known to be very near to 4.803×10^{-10} e.s.u., and hence its mass proves to be 9.02×10^{-28} gm.

Crossed fields of this nature are not of value in electronic practice; in the case of the cathode-ray tube one or the other type of field is used to impose two independent deflexions at right angles upon the electron beam.

*7. Crossed Fields: Electrons emitted from Plate

With electric and magnetic fields crossed orthogonally as before, we now suppose electrons to enter them not in a plane parallel to the deflecting plates but by emission from the inner surface of one of the plates. Such emission may occur simply under the influence of the electric field, or by the photo-electric effect of incident X-rays or light (see section 45).

Let electrons emerge, with zero initial velocity, from the negatively charged plate P_1 and pass towards the positively charged plate P_2 (Fig. 5) across a magnetic field directed into the plane of the diagram. The x -axis is taken normal to the plates, the y -axis in their surface, and the z -axis along the magnetic field, the value of which is accordingly given as H_z . As before, the electric field E_x is given by the quotient of the normal separation of the plates into the potential difference maintained between them.

If emitted in the x - y plane, the motion of a charged particle will be confined to this plane since there is no force acting normally to it; i.e. $d^2z/dt^2 = 0$. In such conditions it is convenient to consider the displacement of the electron in the x and y directions in turn. At right angles to the electric field (y -axis) the only effect is that of the magnetic field in relation to the motion of the electron in the x -direction, and we have

$$m d^2y/dt^2 = e H_x dx/dt. \quad (\text{I.3})$$

The resulting motion in the y -direction is also normal to the magnetic field, giving rise to a magnetic force in the x -direction (Left-hand Rule)

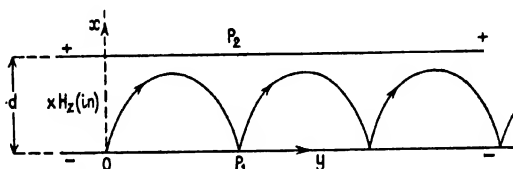


FIG. 5. Cycloidal path of electron in electric and magnetic fields crossed at right angles. H directed into diagram.

which must be added to the electric force, eE_x . The equation of motion along the x -axis is then

$$m d^2x/dt^2 = eE_x - eH_x dy/dt. \quad (\text{I.4})$$

As there is no motion in the z -direction, the position of the electron in the x - y plane can be specified by a complex number of the form

$$u = x + iy.$$

By this substitution, equations I.3 and I.4 can now be combined in a single expression for the motion of the electron. Multiplying I.3 by i and adding to I.4 we have

$$d^2x/dt^2 + id^2y/dt^2 = (e/m)(E_x - H_x dy/dt + iH_x dx/dt).$$

Substituting for u :

$$d^2u/dt^2 = (e/m)(E_x + iH_x du/dt).$$

This is a differential equation of familiar type, with the initial boundary condition that when $t = 0$, $u = 0$, and $du/dt = 0$. The solution is of exponential form:

$$u = (iE_x/H_x)[t - (im/eH_x)(1 - e^{iH_x t/m})].$$

Since

$$e^{i\alpha} = \cos \alpha + i \sin \alpha,$$

we have

$$x + iy = u = (mE_x/H_x^2 e)[iH_x te/m + (1 - \cos H_x te/m - i \sin H_x te/m)].$$

Equating real and imaginary parts in this equation,

$$x = (mE_x/H_x^2 e)(1 - \cos H_x te/m),$$

and

$$y = (mE_x/H_x^2 e)(H_x te/m - \sin H_x te/m).$$

These are the equations of a cycloid generated by a circle of diameter $2E_x m / eH_z^2$, rolling along the line $x = 0$. The maximum possible excursion in the x -direction is thus also given by $2E_x m / eH_z^2$, the distance along the y -axis between the cusps of the cycloid is π times as great, being equal to the circumference of the generating circle, and the average velocity in this direction is given by E_x / H_z .

The cycloidal path (Fig. 5) will only be realized if the separation (d) of the plates is greater than the maximum excursion in the x -direction; the value of the field required to ensure this is $H_z = \sqrt{(2E_x m / ed)}$. For lower values of the field, electrons will arrive at the positively charged plate on their first excursion. By determining the critical value of the magnetic field which sufficed to interrupt the current between the plates, J. J. Thomson was able to measure by this method the value of e/m for the particles emitted from a photo-electric surface, and thus to show that they were electrons. More recently the principle has been applied to the design of electron multipliers, a number of plates at increasing potential along the y -axis being used instead of one continuous plane; the field distribution is thereby distorted, but the electron paths are still to a first approximation cycloidal (see p. 170). Similar conditions obtain also in the magnetron thermionic valve (section 65).

8. Refraction of Electron Beams

The types of transverse fields so far discussed are employed for controlled deflexion of electron beams and not for concentrating them. The conditions are relatively simple, and the electron trajectories can be exactly calculated. In electron lenses, on the other hand, the field distribution cannot normally be expressed in terms of an exact function, nor is it possible to deduce mathematically the electron paths even when the field has been experimentally mapped; it is usually necessary to resort to point-to-point plotting. It is essential in these circumstances to introduce the idea of refraction of an electron beam at a potential boundary; this is an alternative and more useful formulation of the elementary rule, already illustrated in dealing with transverse fields, that electrons moving in an electric field tend towards the direction of the lines of force. We now transfer attention to the equipotential lines (or surfaces) and discuss the deflexion of electrons in terms of successive stages of refraction in passing from one supposed equipotential region to the next. The field will, in reality, always vary smoothly between its boundary values without discontinuities, and the insertion of equipotential lines represents an arbitrary subdivision for convenience of treatment. A medium of continuously varying refractive index is not easy to discuss, until it is replaced by a series of layers of regularly increasing 'density', as in treating refraction in the atmosphere.

The space between any charged electrodes is therefore represented as covered by a set of equipotential surfaces whose density of distribution may be adjusted to the accuracy of treatment required; at all points the lines

of force will be at right angles to them. The problem then is to find the path taken through this system by an electron of given velocity and given initial direction. The deviation suffered at each successive equipotential is evaluated and a point-to-point plot made on the assumption of rectilinear propagation between them, the accuracy increasing with the degree of subdivision of the field.

Suppose an electron of velocity v_0 to approach an equipotential specified by the constant potentials V_0 and V_1 on either side of the surface (Fig. 6). The angles of incidence i and refraction r are defined as in optics with respect to the perpendicular at the point of incidence, O . The energy of the electron measured in kinetic and electrical terms must be the same,

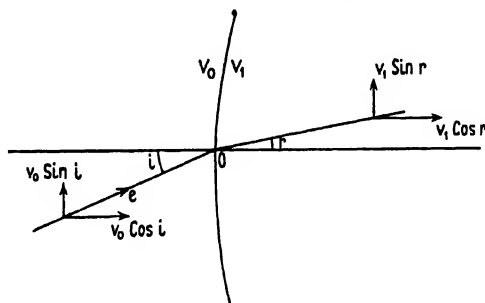


FIG. 6. Refraction of an electron beam at a potential discontinuity ($V_1 > V_0$).

assuming it accelerated from rest through the potential difference V_0 ; that is to say,

$$eV_0 = \frac{1}{2}mv_0^2, \quad \text{or} \quad v \propto \sqrt{V}. \quad (\text{I.5})$$

In crossing the surface the tangential velocity of the electron will be unchanged, there being no electric force in this direction, but the normal component will instantaneously increase from $v_0 \cos i$ to a larger value $v_1 \cos r$ in the region of higher potential. (In practice, the transition will not be abrupt, as assumed here, but will take place over a finite region of varying potential with a finite acceleration.) Then, since

$$v_0 \sin i = v_1 \sin r,$$

$$\sin i / \sin r = v_1 / v_0 = \sqrt{(V_1 / V_0)} = \frac{N_1}{N_0}, \quad (\text{I.6})$$

where N_0 and N_1 may be termed the refractive indices of the respective spaces. This ratio is constant for all incident electrons, and a law equivalent to Snell's law of optical refraction is found to be obeyed. The process is in every way similar to the refraction of a corpuscle in Newtonian optics at the boundary between two media, whose refractive indices now depend on the potentials applied. For all practical purposes we set the refractive index N of a potential space equal to the root of the prevailing voltage. In contrast to wave optics, therefore, an electron increases in

velocity on entering a medium of high 'density' and is refracted towards the normal. In passing into a region of lower voltage the reverse is the case and, when the local potential falls below that of the origin of the electron, reflection will occur.

Practical electron optics thus becomes a matter of determining the configuration of the equipotentials of a given field and of plotting step by step the paths of electrons through it, as is indicated in some detail in the next chapter. With magnetic fields this procedure becomes more complex but fundamental analysis proves to be rather easier, with the result that the plotting of trajectories is only necessary in special cases. Here it will be enough to say that, by proper design of the electrodes or pole-pieces (as the case may be), fields may be set up that act upon electrons in an analogous manner to prisms, mirrors, and lenses in optics. Indeed, the transverse electric (or magnetic) field already discussed forms a simple prism for electrons, since the resulting deflexion varies in magnitude with the velocity of the beam, corresponding to the dependence of deviation on wave-length in the case of light. The production of a lens is then the problem of producing a field in which the deviation of an electron increases in proportion to its distance from the axis; this is not so simple of solution as in light optics.

In light optics the density and boundary shape of the material media are the variables, whereas in electron optics we are concerned with the much less obvious equipotential systems; the important factors now are the shape and separation of electrode surfaces and the ratio of voltages applied to them. For the most part, we shall limit our considerations to paraxial beams moving in fields of circular symmetry, in which the same paths are followed in any plane containing the axis. The charges giving rise to the fields will normally be supposed confined to the electrode surfaces. In a later section, however, the importance in many circumstances of distributed space charge will be indicated (section 44); one or two applications involving cylindrical fields will also be described (Chapter IX).

9. *Limitations of the Optical Analogy*

Although the analogy with optics is of the greatest value in reaching an understanding of electronic phenomena, it must be constantly borne in mind that it has very definite limitations. The most important difference has already been mentioned: that in electrical and magnetic fields there are no sharp changes in refractive index but a continuous variation, and only as a matter of convenience is the inter-electrode space imagined as divided up by equipotential lines. This becomes an even more complicated matter in magnetic lenses, where the refracting field proves to be anisotropic and deviation depends on the *direction* of the beam as well as on its velocity. Comparable cases are met with in optics, but only in advanced work.

A second fundamental difference is the lack of different media in electronics, in contrast with the diversity of materials with varying optical

properties. In optics it is a comparatively simple matter to select the right type of glass, quartz, or other transparent substance for use in a particular case or for building up a corrected lens to a given specification. In electronics we have only the electric and the magnetic field: however much we vary the shape of electrodes and their potentials, the basic interaction of electron and field remains the same. Hence it is impossible to devise one field configuration to correct completely the aberrations of another. Worse still, the aberrations of magnetic fields are in the main of the same sign as those of the electric field so that correction cannot be achieved by combining them. Indeed, the only possibility of full correction would appear to be offered by controlled use of space charge, a difficult technique in which practically no work has yet been done.

Fundamentally for the same reasons it is only possible to construct electrostatic lenses of concavo-convex form (or the converse), including the limiting plano-types: that is to say, meniscus lenses in which the refraction in the two halves is in opposite senses. The convergent effect is always found to predominate, except in one electrostatic type of weak negative power. Biconvex or biconcave systems have not yet been discovered; the symmetrical three-electrode lens (section 26) is equivalent to a combination of two meniscus lenses.

It must be remembered also that an electron path is rarely as free as is the case with light. For the latter, propagation is normally rectilinear and it is the exception for deviation to be introduced, since it occurs only at the boundary between media or by variation in density of a given medium. In electronics there are almost always electric or magnetic fields acting at all points of the electron path, which is, therefore, usually curved and with a continuously varying radius. The path will only be rectilinear where the field lines happen to be straight and the electron is projected along them, or after it has emerged from a field and continues under its acquired momentum in field-free space. Such a region in any case will not be extensive, owing to the physical limitation of having to maintain the experimental system *in vacuo*, on account of the rapid absorption and scattering of electrons by gases. The curvature of the path carries with it, of course, possibilities not present in light optics; these have been exploited in many practical devices (cyclotron; electron multipliers).

On the other hand, both electric and magnetic fields have the great advantage of flexibility. Refractive indices can be varied at will by simple control of voltage or current in a way that is almost impossible in optics. This makes the experimental investigation of electronic systems a comparatively rapid matter, and is the reason why so many problems have been solved empirically whilst theoretical treatment has developed slowly. In addition, the range of refractive indices available by variation of the field is enormous, in comparison with optical media. The maximum lens voltage so far used is in the neighbourhood of 300,000 V., giving an equivalent refractive index (over-all) of about 550 as compared with an upper limit of about 2.5 in optics. In an ordinary cathode-ray tube the possible

range is from 1 to 50 or 60. As a consequence there is normally no necessity to have interchangeable lenses of various powers, as in optics.

Furthermore, by applying pulsating fields it is possible to achieve a type of focusing unknown in optics, whereby electrons from an initially uniform beam are 'bunched' into pulses of large charge density separated by spaces of almost zero current. This so-called 'phase-' or 'space-time' focusing will be mentioned in more detail in Chapter X.

10. Beam Parameters

Before proceeding farther it will be helpful to explain the different methods in use for specifying the velocity of an electron beam. In the case of electric fields this is usually done by reference to its energy, and in magnetic fields to its momentum, these being the practically measurable quantities in the respective cases.

The energy of an electron moving with velocity v is frequently expressed in terms of the voltage V that would accelerate it from rest to this velocity. One electron-volt (e.V.) is, therefore, the energy that an electron acquires when accelerated by a potential difference of one volt. The energy in electron volts is directly related to the kinetic energy in ergs:

$$\frac{1}{2}mv^2 = \frac{Ve_s}{300} = Ve_m 10^8, \quad (\text{I.7})$$

using the conversion factors into electrostatic and electromagnetic units respectively. When it is stated simply that $v = \sqrt{2Ve/m}$, V will be understood to be in absolute units; e is similarly measured, unless otherwise stated, in electromagnetic units. As an example, an electron accelerated from rest through a potential difference of 1,000 V. will emerge with an energy of $10^8 \times 10^8 \times 1.601 \times 10^{-20} = 1.601 \times 10^{-9}$ ergs and a velocity of 1.87×10^9 cm. sec.⁻¹ One e.V. is thus equivalent to 1.601×10^{-12} ergs. If particles of other kinds are in question, regard must be had to both their mass and charge: under the same accelerating voltage a proton would reach a velocity of 4.38×10^7 cm. sec.⁻¹, its mass being 1,839 times that of the electron and its charge the same, whilst an alpha particle with twice the charge and 3.97 times the mass of a proton would reach 3.11×10^7 cm. sec.⁻¹

At higher accelerating potentials account must be taken of the relativistic increase in mass over the rest-mass m_0 , given by

$$m = \frac{m_0}{\sqrt{1-v^2/c^2}},$$

where c is the velocity of light; v/c is usually denoted by β . The kinetic energy is then given as

$$\begin{aligned} (m-m_0)c^2 &= m_0 c^2 \left\{ \frac{1}{\sqrt{1-\beta^2}} - 1 \right\}, \\ &= \frac{1}{2}m_0 v^2 (1 + \frac{3}{2}\beta^2 + \dots), \end{aligned}$$

on binomial expansion. The resulting correction to the velocity calculated from the classical expression (I.5) is less than $\frac{1}{2}$ per cent. up to 3,000 V.

and less than 1 per cent. at 7,000 V. Practically, the correction is only of importance in high-voltage cathode-ray tubes, electron diffraction apparatus, and electron microscopes, where voltages of over 10 kV. are in question. At 260 kV. an electron beam will have 0.75 of the velocity of light, whereas in elementary theory it should be travelling faster than light under this accelerating potential.

In magnetic fields the radius of the deflected electron path is the experimental quantity, and its momentum is quoted in terms of equation I.1: $Hr = mv/e$, with the relativistic correction for mass at high velocities. If e is in electromagnetic units, then Hr will be in oersted-centimetres. Thus the (Hr) value for a beam of 1,000 e.V. electrons will be 107, and for a 260 e.kV. beam will be 1,925 oersted-cm.

The equivalent wave-length (λ) of an electron is given (cf. section 58) by de Broglie's relation

$$h/\lambda = mv = eHr, \quad (\text{I.8})$$

where h is Planck's constant (6.624×10^{-27} erg-sec.). Inserting the values of the constants, and neglecting relativity effects, this takes the useful practical form

$$\lambda = \sqrt{(150/V)}, \quad (\text{I.9})$$

giving λ in angstrom units when V is in volts. Thus the 1,000-V. beam will have an equivalent wave-length of 0.387 Å. and the 260-kV. beam a value of 0.0213 Å., in each case much smaller than atomic dimensions.

It follows that inhomogeneities of velocity in an electron beam are exactly equivalent to variety of wave-length in a light beam and will result in dispersion effects in electric or magnetic fields. The analogue to a 'monochromatic beam' will be a beam in which all electrons have the same velocity; near-monochromatic sources are the rule in electronics, as fluctuations in accelerating voltage make it difficult to obtain a truly uniform beam otherwise than by controlled dispersion.

FURTHER READING

J. J. Thomson, *Phil. Mag.* **44**, 293, 1897.

— *ibid.* **48**, 547, 1899.

$$1 \text{ keV} - 107 \quad .387 \text{ Å}^{\circ}$$

$$260 \text{ keV} - 1925 \quad .0213 \text{ Å}^{\circ}$$

$$\lambda = \sqrt{150/V} \text{ Angstroms} \quad \downarrow \text{volts}$$

CHAPTER II

THE ELECTROSTATIC FIELD

11. General

THE fundamental problem in an electron optical system is to devise an arrangement of electrodes and applied potentials which will cause electrons to travel along predetermined paths, and usually to a sharp focus in a point or plane. It was solved in a certain degree by empirical methods before a basis for theoretical procedure was set out by Busch in

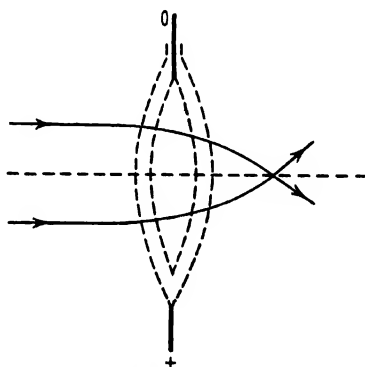


FIG. 7. Mesh electron lens.

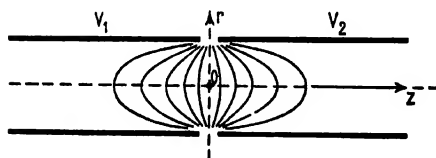


FIG. 8. Electron lens formed between two concentric cylinders.

1926. A number of experimenters had found that an electrode assembly similar to Fig. 27 gave good focusing of the electrons emitted from a cathode C . G is a negatively biased shield and A an anode, positively charged with respect to the cathode; in practice A was usually earthed and C maintained at a high negative potential.

Busch discussed the far-reaching analogy between light and electron optics and succeeded in evaluating the trajectories of electrons, by rather laborious methods, in both electrical and magnetic fields. The obvious arrangement of convex equipotentials, analogous to the glass lens, was at once tried by applying a potential difference between two shaped

graticules (Fig. 7). It was found to give a focus and to obey to a considerable degree of accuracy the simple lens formula relating object, image, and focal distance ($1/u + 1/v = 1/f$). The gauze frameworks, however, are found to interfere with the electron beams both by direct collision and by distortion of the 'spherical' field in the neighbourhood of the wires; they are also unsuited to the application of high voltages, and hence their effective refractive index is limited. In practice they have been superseded by tube lenses, with or without the addition of diaphragms. Fig. 8 shows the distribution of equipotentials, in any plane containing the axis, between two cylindrical electrodes maintained at different potentials V_1 and V_2 . The similarity between these 'refracting surfaces' and the spherical form of optical lenses is evident and, in fact, the system behaves in every way as a lens for electrons. The deviation produced is closely proportional to the distance of an electron from the axis and paraxial beams are brought to a sharp focus; marginal rays suffer a degree of spherical aberration dependent on the construction of the lens tubes and the ratio of the applied voltages. The different arrangements that have been found practically useful will be described in the next chapter. In this chapter we deal with the determination of field distributions and with the methods that have been devised for tracing experimentally the paths of electrons through them.

In theory it should be possible to calculate the solution to any electrostatic focusing problem; that is, given the origin and end of an electron path, to find the shape of the electrodes and the applied potentials necessary to yield the path. In the present state of mathematical technique this proves to be impossible except in the simplest cases, the solutions to the appropriate differential equations being lacking. The converse problem is only slightly more amenable to treatment: given an electrode configuration (i.e. the boundary conditions), to find the equipotential surfaces and hence the paths of electrons through them. The strict solution of a few simple cases will be given here; in practice, methods of progressive approximation are generally used, both in calculation and in experiment.

12. Determination of the Potential Field

Given an electrode system, from which space charge is absent, the essential first stage is to determine the position of the equipotential surfaces in the space between them. Formally this is a matter of finding a solution to the general equation of Laplace,

$$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0, \quad (\text{II.1})$$

with known boundary conditions, V being the potential at any point (xyz) in the space under investigation, as described in rectangular coordinates.

The field in the x -direction is then given by a term $E_x = -\frac{\partial}{\partial x} V(xyz)$, with

equivalent expressions, E_y and E_z , for the field in the y - and z -directions. It follows also that

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0, \quad (\text{II.2})$$

which is essentially an equation of continuity and is equivalent to the statement that as many lines of force enter any volume of free space as leave it. Such an expression is termed the divergence of the vector quantity concerned, written (in this case) as $\text{div } E$ or ∇E ; the field being the rate of change of potential, $\nabla^2 V$ is then the divergence of the gradient of the potential ($\text{div grad } V$).

When space charge is present, of density ρ per unit volume, we have the more general Poisson's equation

$$\nabla E = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 4\pi\rho,$$

indicating a net outward flux of $4\pi\rho$ lines from each unit volume.

Complete solution of these equations would give the potential $V(xyz)$ at any point and hence make possible the exact delineation of equipotential surfaces. An analytic solution is possible only in the geometrically most simple cases, and in the absence of space charge, where the problem is essentially the same as the determination of the capacity of a given arrangement of conductors considered as a condenser: for parallel infinite planes, concentric spheres, sphere and plane, and the analogous cylindrical systems.

(a) *Parallel Plane Electrodes.* Considering the planes (Fig. 9) to be of infinite extent and at potentials V_0 and V_n , the boundary conditions are $E_y = E_z = 0$ in both electrode surfaces, i.e. when $x = 0$ and when $x = d$, the separation of the planes (the x -direction being normal to the planes). The solution of Laplace's equation then has the simple form

$$V = V_0 + Ax.$$

Then the field strength in the x -direction will be

$$E_x = -\frac{\partial V}{\partial x} = -A = -\frac{V_n - V_0}{d},$$

and in the two dimensions parallel to the planes will be

$$E_y = E_z = -\frac{\partial V}{\partial y} = 0.$$

Hence the equipotential surfaces will be equally spaced planes parallel to the electrodes. Electrons emitted normally to one of the latter will follow

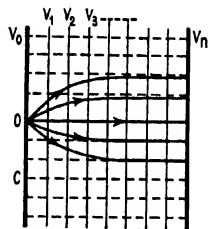


FIG. 9. Field and electron trajectories between plane parallel electrodes. (Lines of force shown dotted.)

straight-line paths to the opposite electrode, that is to say, will travel along the lines of force at right angles to the equipotential surfaces. The equation of motion is

$$\frac{d^2x}{dt^2} = -\frac{eE_x}{m},$$

and hence at any instant t , for an electron starting from rest at $x = 0$ when $t = 0$, we have $v = -(eE_x/m)t$, and $x = -\frac{1}{2}(eE_x/m)t^2$. The complete path is therefore known, and the transit time may at once be evaluated.

For the general case, when an electron is emitted at an angle to the lines of force, its path will be parabolic and finally asymptotic to a line of force as indicated in the diagram. This is approximately the state of affairs when a plane cathode is arranged parallel to a disk anode which has a small aperture at its centre.

(b) *Concentric Spheres.* The solution of Laplace's equation in this case will be $V = -A/r + V_0$, in terms of spherical coordinates, so that

$$E_r = -\frac{\partial V}{\partial r} = -A/r^2.$$

At right angles to the radius at any point the field strength will be zero, and the equipotentials will be concentric spheres. Electrons emitted normally from one surface will travel along the radial lines of force to the other surface (Fig. 10). Only in one practical case, the so-called 'point-projector microscope', is this condition approximately met (see section 49).

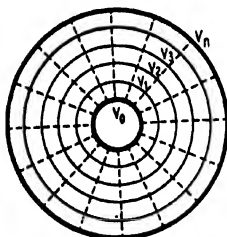


FIG. 10. Field between two concentric spheres. (Lines of force shown dotted.)

(c) *Concentric Cylinders.* In this case the solution is $V = A \log_e r + V_0$; hence the radial field strength is $E_r = -A/r$. The equipotential surfaces are concentric cylinders and the lines of force their radii. Such a system is approximately that obtaining in a triode thermionic valve with a rod cathode and cylindrical grid and anode. In the absence of space charge (i.e. with positive grid potential) the electrons will then follow nearly radial paths from filament to anode.

*(d) *Coaxial Cylinders.* The arrangement of two cylinders on a common axis and at different potentials, constituting as it does the simplest electron lens (Fig. 8), is of especial interest. It does not yield a simple solution, as in the cases just discussed; a number of attempts have been made by the method of series expansion, and the most successful treatment will be given here.

In cylindrical coordinates (r, z, θ) Laplace's equation becomes

$$\nabla^2 V = \frac{\partial^2 V}{\partial z^2} + \frac{1}{r} \frac{\partial V}{\partial r} + \frac{\partial^2 V}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2 V}{\partial \theta^2} = 0.$$

The last term will be zero in lenses of rotational symmetry, as in the present case, and we may write

$$\frac{\partial^2 V}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial V}{\partial r} \right) = 0. \quad (\text{II.3})$$

Assuming that the gap between the two cylinders is infinitely small, it is legitimate to consider the potential as a linear combination of functions, in which the variables are separated, writing

$$V(r, z) = V_1 + V_2 + \dots V_k + \dots,$$

where

$$V_k(r, z) = F_k(z) G_k(r);$$

that is to say, when r is constant the potential is a function of z only, and when z is constant it is a function of r only. On substitution in equation II.3 the result is two ordinary differential equations:

$$\frac{1}{F} \frac{d^2 F}{dz^2} = k^2, \quad \frac{1}{rG} \frac{d}{dr} \left(r \frac{dG}{dr} \right) = -k^2.$$

The general solution of the first is of exponential form:

$$F_k(z) = ae^{iks} + be^{-iks}.$$

The second may be rewritten as

$$\frac{\partial^2 G}{\partial r^2} + \frac{1}{r} \frac{\partial G}{\partial r} + k^2 G = 0,$$

which is a Bessel's equation; its general solution involves Bessel (J_0) and Neumann (Y_0) functions:

$$G_k(z) = AJ_0(ikr) + BY_0(ikr).$$

The solution of Laplace's equation may then be written

$$V(r, z) = \int A(k) F(zk) G(rk) dk,$$

integrating over the entire complex domain, since k does not necessarily have discrete values. The coefficients $A(k)$ are now found from the given boundary conditions by the usual method for Fourier coefficients. Since the potential must remain finite as z increases, all terms with complex k values are eliminated from both $F_k(z)$ and $G_k(r)$. Further, as the potential must be finite along the axis, the coefficients of the Neumann function prove to be zero. Since $V - \frac{V_1 + V_2}{2}$ (where $\frac{V_1 + V_2}{2}$ is the potential at the origin) is an odd function of z , it also follows that the only trigonometrical functions to be considered are sines. The solution finally has the form

$$V(r, z) = \int_0^\infty B(k) J_0(ikr) \sin kz dk + \frac{V_1 + V_2}{2}. \quad (\text{II.4})$$

We now take the radius of the cylinders as the unit of length, and find the coefficients $B(k)$ using the boundary conditions for $r = 1$:

$$V(1, z) = V_1, \text{ for } z < 0, \text{ and } V(1, z) = V_2, \text{ for } z > 0.$$

We get

$$B(k)J_0(ik) = \frac{1}{\pi} \left\{ \int_{-\infty}^0 -\frac{V_2-V_1}{2} \sin kz \, dz + \int_0^{\infty} \frac{V_2-V_1}{2} \sin kz \, dz \right\}$$

and

$$B(k) = \frac{V_2-V_1}{\pi k J_0(ik)} (1 - \lim_{z \rightarrow \infty} \cos kz).$$

Equation II.4 then becomes

$$V(r, z) = \frac{1}{\pi} \int_0^{\infty} \frac{V_2-V_1}{k J_0(ik)} J_0(ikr) \sin(kz) \, dk + \frac{V_1+V_2}{2}.$$

As the potential at (r, z) is related to the axial value (cf. equation II.12), it is enough to know the potential along the axis and its first and second derivatives in order to plot electron trajectories through the lens. In this case we have

$$\left. \begin{aligned} V(0, z) &= \frac{1}{\pi} \int_0^{\infty} (V_2-V_1) \frac{\sin kz}{k J_0(ik)} \, dk + \frac{V_1+V_2}{2}, \\ V'(0, z) &= \frac{1}{\pi} \int_0^{\infty} (V_2-V_1) \frac{\cos kz}{J_0(ik)} \, dk, \\ V''(0, z) &= -\frac{1}{\pi} \int_0^{\infty} (V_2-V_1) \frac{k \sin kz}{J_0(ik)} \, dk, \end{aligned} \right\} \quad (\text{II.5})$$

where primes indicate derivatives with respect to z . Unfortunately, however, it is still impossible to solve these integrals analytically; in practice, they are evaluated by quadrature in order to get numerical values for the axial potential. Having done this, it is possible to plot the paths of representative electrons through the lens, determine the principal planes, and find the focal lengths in the manner described later. The method is obviously extremely laborious, and usually resort is had to direct experimental investigation of the potential distribution with the aid of an electrolytic tank, as described below.

When the gap between the cylinders is not negligible the problem becomes still more difficult, although Bertram has obtained solutions for special forms of the potential distribution in such a lens.

(e) *Conformal Mapping.* In cases where the variation of potential is confined to two dimensions, the whole field being described by moving a plane parallel to itself, a simpler solution is sometimes possible; this condition obtains in the space between deflecting plates and in most types of electron multiplier. Laplace's equation then reduces to

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0. \quad (\text{II.6})$$

Since lines of force are by definition orthogonal to equipotential lines, the

system can be represented electrostatically by a suitable complex function of the form

$$\phi + i\psi = f(x + iy),$$

it being known that such functions satisfy the Laplace condition. Here ϕ represents the potential and ψ the equation of a line of force, or vice versa, since they are conjugate quantities.

Straightforward solution of this general equation is again impossible except with the simplest boundary conditions. The usual procedure is, in fact, to work backwards by taking some tractable function of the complex variable and plotting the families of curves that result for the ϕ and ψ functions. Since any curve of either function can be taken as a fixed boundary, it may then be possible to fit some of the configurations obtained to practically significant arrangements of electrodes. As an illustration we may examine the simple form

$$\phi + i\psi = (x + iy)^2 = x^2 + 2ixy - y^2.$$

On separating real and imaginary quantities, we have $\phi = x^2 - y^2$ and $\psi = 2xy$. The equipotential curves will then be rectangular hyperbolas

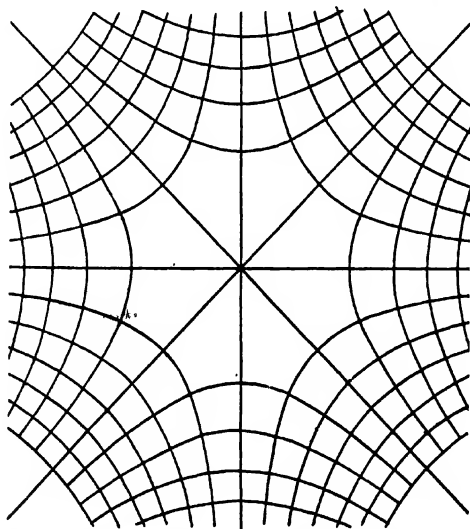


FIG. 11. Equipotential and field lines for the two-dimensional distribution $\phi + i\psi = (x + iy)^2$.

with the x - and y -axes as their axes, and the lines of force will be rectangular hyperbolas with these lines as asymptotes (or vice versa). The resulting families of lines are shown in Fig. 11. From this system we have at once the solution to the field between two charged planes intersecting at right angles and inclined at 45° to the x - and y -axes. Or, by taking corresponding lines of force in each quadrant, we have the configuration of the field in the space formed by the junction of four channels.

Other systems to which comparatively simple conjugate functions can be fitted are the single conducting plane and two parallel cylindrical wires. It has long been a standard method in electrostatic theory for finding field distributions; unfortunately it has no general validity, and is distinguished more for its elegance than for any great practical importance.

13. Direct Calculation of Field Distribution by the Relaxation Method

Most of the electrode systems employed for focusing electrons are more complex than the simple forms considered above for which a rigorous analytic solution is possible. Even in the case of two coaxial cylinders of equal diameter the solution of Laplace's equation led to a result which is

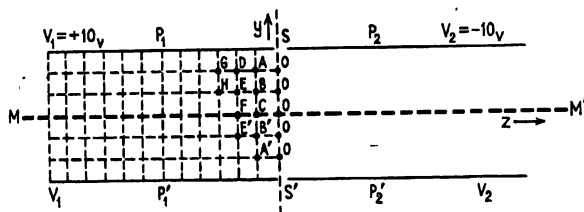


FIG. 12. Formation of grid between plane parallel plates for calculation of field distribution.

difficult to apply in practice; when the gap between them is finite, or when they are of different diameters, then analysis fails and reliance has hitherto been placed on practical measurements of field distribution by one of the methods described later. Very recently, however, Motz has shown that the relaxation method due to Southwell can be readily used to calculate the field in all cases where definite boundary conditions exist. Starting with these known conditions, usually the form and potentials of the electrodes, a method of successive approximation allows the field distribution to be calculated to any specified degree of accuracy by arithmetical processes only. The case of two concentric cylinders of unequal diameter is fully described by Motz and Klanfer, but the principles of the method are better explained by taking a simpler example, for which an analytical solution is available.

Suppose a slot SS of negligible width to be cut in two infinite parallel plates, thus dividing them into two symmetrical pairs, P_1P_1' and P_2P_2' , to which different potentials V_1 and V_2 are applied (Fig. 12); the field will be symmetrical also about the median plane MM' , and therefore calculation need only be made for one quarter of it. According to the degree of accuracy desired, the space between the plates is divided into a grid of greater or less mesh-length. For ease of treatment we here take a large mesh and convenient values for the potentials: $V_1 = +10$ V., $V_2 = -10$ V.; the potential in the gap-plane SS will then be zero. In this system the potential is independent of the x -coordinate and we need only find the field distribution in one y - z plane; that is, at the lattice points A, B, C, D, \dots

only. Laplace's equation for the two-dimensional case, as already stated (II.6), is of the form $\partial^2 V / \partial y^2 + \partial^2 V / \partial z^2 = 0$.

A solution may be obtained in the form of a series of even derivatives of the potential at a given point, the odd powers vanishing on account of symmetry. From this expression it is shown that the potential V_0 at any point O is given by the arithmetic mean of the values at its immediate neighbours P, Q, R, S , in the y - and z -directions, if their spacing is small enough to allow terms of higher order than the second to be dropped from the expansion. Thus

$$(V_p + V_q + V_r + V_s) - 4V_0 = 0 \quad (\text{II.7})$$

to the second order of approximation.

The method consists in assuming initially an arbitrary distribution of

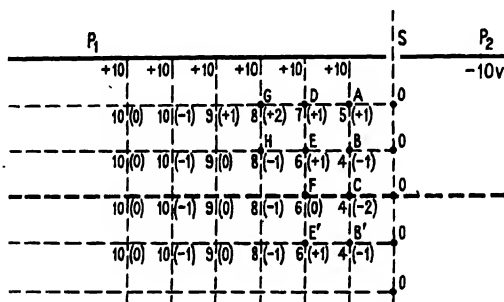


Fig. 13. First approximation to potential distribution.

potential over the lattice points and then adjusting the values successively, until the residual differences obtained by the application of II.7 vanish or are reduced below an assigned limit at every point. It is usually convenient to start from a point such as A next to two known potential boundaries, and with the assumption that the potential has a uniform value at all lattice points: in this case, say, $+10$ V. The first application of the criterion II.7 shows that the residuals are zero at all points save those such as A , which are near the zero equipotential of the gap-plane (Fig. 12). We try the effect on V_a of reducing its assigned potential by 1 V., and note also the effect on neighbouring points: its own (negative) residual is reduced by 4 and those of B and D are changed by -1 . We proceed to adjust the values at B and C similarly, noting that a change in V_b doubly affects the potential at C since the same change must be made at the symmetrical point B' ; this is true for all points next the axis. The same process is applied in turn to points D, E, F , etc., recording always the effect of a change at any point upon its immediate neighbours, and returning to make adjustments in a potential whenever its residual exceeds an assigned value, which may be ± 2 at a first trial. In this way the approximate potential distribution within this sector may be rapidly determined, as in Fig. 13, where the final residuals are shown in brackets.

For higher accuracy the potential at every point in the first solution is now multiplied by 10, as are also the corresponding residuals, and the process of trial repeated. The distribution is found to be that shown in Fig. 14; comparison with the analytical solution shows the error now to be about 2 per cent. In regions where the potential gradient, and therefore also the effect on an electron, is greatest, it is desirable to explore the field with higher accuracy: the lattice interval is halved and the process of trial repeated. The student should check for himself that the potentials

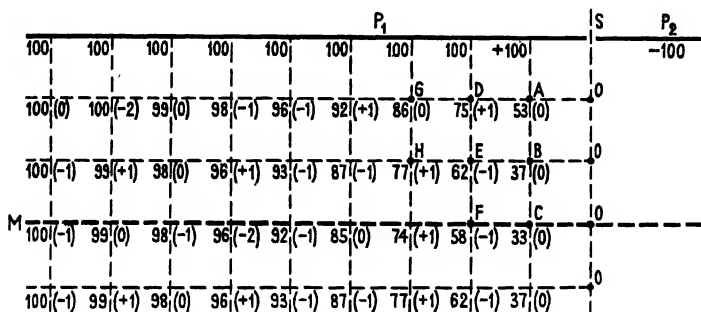


Fig. 14. Second approximation to potential distribution.

of points (*A*, *B*, *C*) near the gap are rarely changed, except in the residuals, as a result of this process. In this way a complete map of the field may be obtained to any required degree of accuracy, if the grid is made smaller and smaller in mesh. The final refinement is to observe the effect of the neglected fourth-order terms, using an expression corresponding to II.7; the new values of the residuals thus obtained are then reduced by further adjustment of the potentials, using the second-order equation. The greater the refinement of the investigation, the greater is the labour involved, and a calculating machine is then employed; but a general picture of the form of the equipotentials may be gained very quickly by simple arithmetic, employing II.7 only.

The method is readily extended to three dimensions and therefore to practical electron lenses. It is then more convenient to use cylindrical rather than rectangular coordinates, the corresponding condition to II.7 being

$$V_P(1-a/2r) + V_Q + V_R(1+a/2r) + V_S - 4V_0 = 0,$$

where a is the mesh-length and r the radial distance of O , Q and S from the axis; the radial distance of P is $r-a$ and that of R is $r+a$. The method may also be employed in the reverse sense: to determine the necessary form and potential of the bounding electrodes required to produce a given axial potential distribution. A linear rise in axial potential between two regions of zero field, for instance, may be quickly shown to require the complicated electrode form of Fig. 15, in which alternate ring segments are at positive and negative potential.† Scherzer has obtained

† Motz (unpublished).

a similar result analytically (cf. p. 128). The relaxation method may be applied also to magnetic fields, and to combined electrostatic and magnetic fields, and thus constitutes one of the most powerful means yet devised for the investigation of electron optical systems.

14. *Experimental Determination of the Potential Field*

An analytic solution for the field distribution is only obtainable in such simple boundary conditions as have been discussed above; the case of coaxial cylinders of equal radius is already difficult to handle mathematically. In practice, the electrode arrangements are almost always much more complex, the cylinders being fitted with diaphragms, having differing radii, and sometimes being more than two in number. It is then

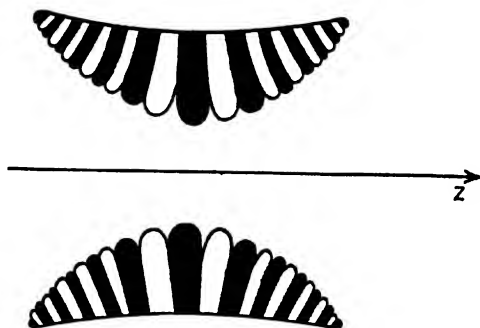


FIG. 15. Form of electrodes required to produce a linearly increasing axial potential.

helpful to make use of the analogy between electrostatic fields and those that can be established electrolytically or gravitationally. By this means the distribution of the field can be experimentally explored to a high degree of accuracy.

The complete analogy between the electrostatic field around electrodes in vacuum and that existing around electrodes of similar shape and relative potentials immersed in a conducting solution has been appreciated and utilized for many years. It was first described in 1913 by Fortescue and Farnsworth, and has recently been applied by many investigators to the solution of electron optical problems. The basis of the analogy is the equivalence of the flux of lines of electric force with the flow of current in an electrolyte. The field strength at any point in the latter, being an ohmic conductor, is directly proportional to the current density, which in turn obeys an equation of continuity: as much current leaves as enters any volume element in the solution. The divergence is zero, and Laplace's equation is satisfied for the potential distribution around the electrodes. The potential at any point between them is then proportional to the potential at the corresponding point in the electron optical system.

The practical procedure is to set up a large-scale model of the electrodes in an electrolytic 'trough', as it is called, to charge them from a battery

or generator to the same relative potentials as in the original, and to investigate the potential distribution with a conducting probe. The trough (*T*, Fig. 16) should be made of glass or impregnated wood to ensure that the equipotentials are perpendicular to the walls, which then do not affect the field to be measured. It must also be large enough in relation to the electrode model to avoid disturbance of the field, say 2 ft. broad by 2 ft. deep and 4 to 6 ft. long, covered on the outside with sheet metal as an electrostatic shield. As very small currents are adequate, ordinary tap-water is usually found to be sufficiently conducting, or a standard copper-sulphate solution of a few milligrams per litre may be preferred. The probe *P* is a short metal wire, preferably platinum, of about 0.02 mm. diameter, connected to a wire or an arm so that it can be lowered into contact with the

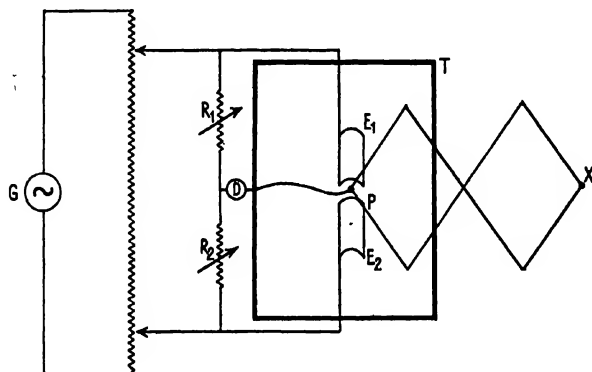


FIG. 16. Electrolytic trough for measuring potential distribution.

surface. Variable potentials are supplied to the electrode models from two tapping-points on a potentiometer, which may consist of two variable resistances of the order of 10,000 ohms each. This potential difference is applied also across two arms R_1 , R_2 of a Wheatstone's bridge, the other two arms being formed between the probe and the two electrodes E_1 and E_2 in the solution. The source of potential *G* should be an A.C. generator rather than a battery, in order to avoid electrolysis and polarization troubles; a valve oscillator working at 400–1,000 cycles per second is frequently used. The detector *D* will then be not a galvanometer but a telephone receiver. When minimum sound is heard, the probe will be at a point whose potential is given by the ratio of the values of R_1 and R_2 .

If the electron optical system has no simple symmetry, then complete scale models of the electrodes must be made and totally immersed in the solution, making the exploration of the space between them a matter of some difficulty. Most lenses, however, possess some degree of symmetry and can be divided by a plane through the axis into two symmetrical portions; across this plane there will thus be no current flow when the system is charged, or, in other words, the equipotentials must be perpen-

dicular to the plane of symmetry. The same condition holds at any insulating surface, such as the base of the tank or the surface of the liquid (since the conductivity of the air above it is negligible). It is, therefore, justifiable to make a model of half the electrode system only, divided longitudinally by a plane pp' through the axis (Fig. 17). The model is placed with this plane in the surface of the solution, and the potential distribution in it then investigated by exploring it with the probe, which just makes contact with the surface. When, as in most cases, the lens has full rotational symmetry there are an infinite number of equivalent planes that could be chosen to bisect the system. It is then enough to take as model not one-half of the original, but only a wedge-shaped segment S formed by two intersecting planes embracing a few degrees of angle (see

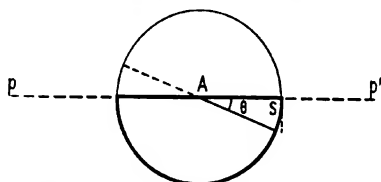


Fig. 17. Cross-section of model of cylindrical electrode.

Fig. 17). Both sides of the wedge must be enclosed by insulating barriers, of course, and this is achieved by using a tank with a sloping bottom of glass on which the model segment rests, its other edge being in the surface of the solution; the thin end of the wedge A , corresponding to the axis of the original system, must then coincide with the line of junction of the liquid surface with the trough bottom. Since the angle θ of the wedge can be made small, and the form of the field near to the electrode surface is usually not required, the curvature of the latter is often neglected and flat strips of metal (bent, if need be, to or from the axis when the electrodes are not of constant radii) are taken, as indicated by the dotted line in Fig. 17. The preparation of models and the trial of various forms is by this means greatly simplified. The method is valuable for obtaining a general picture of the field in the plane of symmetry. On account of surface tension effects, however, it does not give a reliable indication of the field near the axis, which is of first importance for the plotting of electron trajectories.

For accurate investigation, complete half-sections of the electrodes must be employed. If an extensive exploration of the field is required, it is then convenient to set the ratio of the resistances R_1 and R_2 at some simple value, such as 1:9, and to trace with the probe across the surface of the liquid the path for which zero deflexion (or sound) is recorded by the detector; this path will be the corresponding equipotential of the field, in this case at 10 per cent. of the potential difference between the electrodes. If the first were at 500 V. and the second at 3,000 V., the 10 per cent. equipotential line would correspond to a voltage of $500 + (1/10 \times 2,500)$, or 750 V. The ratio of R_1 to R_2 can then be adjusted to correspond to further

equipotentials at any desired intervals, according to the degree of accuracy required. In order to plot the shape of the lines directly, the probe *P* is often fixed to one end of a pantograph linkage with a pencil or stile *X* at the other end, as shown in Fig. 16; as each equipotential is traced out in the surface, the corresponding plot is made on the drawing-paper. In this way the field distribution is rapidly obtained, in the form shown in Fig. 18. This method is preferable to fixing uniform steps of subdivision in hundreds or thousands of volts, since the shape of the field is determined by the shape of the electrodes and will not change if the voltages are altered. A general plot in terms of percentages of the potential difference allows the designer to use it for any other voltage ratio as required.

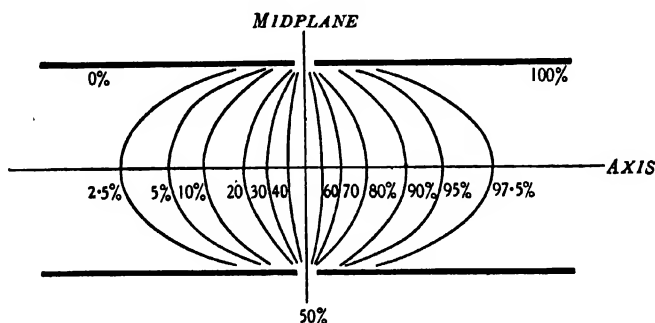


FIG. 18. Potential distribution in the symmetrical two-cylinder lens.

It will be seen that the field, in this case of two identical electrodes, is entirely symmetrical about the median plane. Rotation of the diagram about the axis will give the complete distribution of the field in space. The potential gradient is clearly greatest near the middle of the lens, where the lines are closest together; the 97.5 per cent. equipotential is roughly twice as far from the median plane as the 90 per cent. line. If a closer investigation of any part of the field is necessary, then thin metal sheets bent to the appropriate shape can be inserted in place of any two equipotentials, and the corresponding potentials applied to them. They can be made as large as the trough will allow, and hence a large-scale plot of the field in any part of the lens can be readily conducted.

Frequently it is sufficient to know the axial variation of the field, and not the complete map in a plane of symmetry. It is then only necessary to find the ratio of the resistances needed to balance the circuit for successive positions of the probe along the axis, the intervals depending on the accuracy required. Alternatively a voltmeter can be used to give direct readings. As it draws no current a cathode-ray tube of known deflexion sensitivity is to be preferred for this purpose, one pair of deflecting plates being connected between the probe and one of the electrodes. A curve such as that shown in Fig. 19 is obtained for the simple two-cylinder lens, from which the first derivative V' of the axial potential can

be at once measured (dotted line). For many methods of finding trajectories it is also necessary to know the second derivative V'' , but graphical evaluation from the first derivative is not sufficiently accurate. If, however, a plot of the equipotentials in the neighbourhood of the axis is made and their radius (ρ) measured directly, assuming them to be circular in this plane, then it can be deduced from Laplace's equation that

$$\rho = 2V'/V'',$$

allowing the second derivative to be found with sufficient accuracy. The variation of the two derivatives along the axis is shown in Fig. 19.

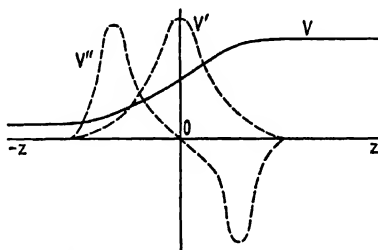


FIG. 19. Variation of axial potential, and of its first and second derivatives (V' , V'') in two-cylinder lens.

15. Motion of Electrons in a Potential Field

Theoretically, once the distribution of the field has been found either by calculation or in the electrolytic trough, it remains to determine the motion of an electron in it from the elementary equations of motion:

$$m d^2x/dt^2 = -eE_x = e dV/dx,$$

$$m d^2y/dt^2 = -eE_y = e dV/dy,$$

$$m d^2z/dt^2 = -eE_z = e dV/dz.$$

Formally it is only necessary to insert values for the potential and eliminate t in order to find the path of the electron. In the case of plane parallel electrodes in crossed electric and magnetic fields (p. 8) it was possible to apply this method because of the simple geometrical conditions and the uniformity of the fields. In practical cases, however, we usually cannot represent the potential distribution by any simple mathematical function, and no analytic solution is possible. The trajectory can only be found by adopting simplifying approximations or by graphical methods. For rough working it is often accurate enough to assume the lens to be composed of a number of lens segments of increasing refractive index, and to plot the electron path from surface to surface using the analogous equation to Snell's law (I.6). At each surface the refractive index for the electron is given by the root of the voltage ratio, $\sqrt{(V_2/V_1)}$. Having drawn the normal at the point of incidence of the electron and measured the angle of incidence (i), the angle of refraction (r) can be calculated and traced on the

diagram, giving the path of the electron (assumed straight) to the next equipotential surface. Such a plotted trajectory is shown in Fig. 20. The accuracy of the method, however, is very low owing to the difficulty of exact measurement of the angles with the normal and of exact construction of the normals themselves. The errors are cumulative, and the image point may be inaccurate by as much as 30 or 40 per cent. It is more usual to use either the type of graphical plotting developed in light optical practice, or an approximate method of calculation from the axial potential and its first and second derivatives.

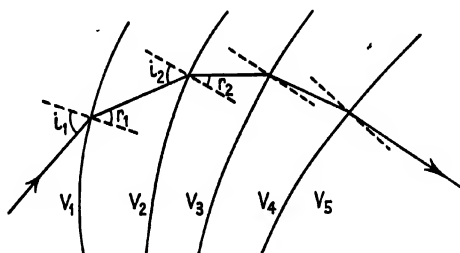


FIG. 20. Electron trajectory plotted from refraction at equipotential surfaces.

16. Trigonometrical Ray-tracing

The method long used in light optics for ray-tracing has been adapted by Klemperer and Wright to electron optical problems. The lens is supposed divided into n sections, bounded by the equipotentials, V_1, V_2, \dots, V_n . It is assumed that the initial direction of the electron is known and that the radius of curvature ρ of the equipotentials at the axis has been measured, assuming them to be of spherical shape in this region. Consider the first equipotential, of radius ρ_1 and centre of curvature C_1 , encountered by an electron starting in a direction making an angle θ_1 with the axis (Fig. 21); let this direction intersect the axis at A and meet the equipotential at X . Drop a perpendicular from C_1 on to AX to meet it at Y . Then C_1X is a radius of the surface and defines the angle of incidence i , which can be evaluated from the simple relations

$$\sin i = C_1Y/\rho_1, \quad \sin \theta_1 = C_1Y/(AE + \rho_1),$$

or

$$\sin i = \frac{u + \rho_1}{\rho_1} \sin \theta_1, \quad (\text{II.8})$$

writing u for the distance of the object point from the refracting surface, as in optics.

The direction of the refracted ray is given by Snell's law:

$$N_1 \sin i = N_2 \sin r,$$

where N_1 and N_2 are the refractive indices on the two sides of the equipotential. Let this direction XZ be produced backwards to cut the axis at

B , giving the image point and the image distance v . By geometry we then have

$$r = i - \theta_1 + \theta_2,$$

whence the direction θ_2 of the refracted ray with respect to the axis can be found; for this a similar expression to II.8 obviously exists:

$$\frac{v + \rho_1}{\rho_1} = \frac{\sin r}{\sin \theta_2}.$$

The image distance v can thus be calculated, and all the data necessary are now available for repeating the process at the next equipotential surface.

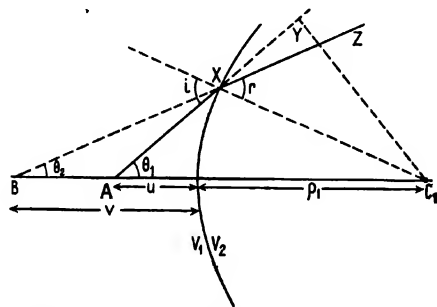


FIG. 21. Trigonometrical ray-tracing at spherical equipotentials.

In this way it is possible to follow the path of an electron through a field that has been mapped in equipotentials in the electrolytic trough. It is convenient for calculation, however, to divide the field by surfaces of equal refractive power and not by regular steps of 5 or 10 per cent. of the total potential difference $V_n - V_1$. That is to say, the ratio of the root of the voltages at each step should be constant, since (by I.6) the refractive index depends on this ratio. The field is therefore divided by a number (n) of equipotentials, depending on the degree of accuracy required, so that

$$N_1/N_2 = \sqrt{(V_1/V_2)} = N_n/N_{(n+1)} = n\sqrt{(V_1/V_n)}.$$

The procedure becomes more laborious as n is increased, and it is usually accurate enough to take $n = 20$. The position of these equipotentials can be readily found from the data obtained with the electrolytic trough by plotting $\log V$ against z , the axial coordinate, and dividing off $\log V$ into n equal steps.

For paraxial rays, for which the assumption of a spherical form for the equipotentials is allowable, the trigonometrical method gives an accuracy of a few per cent. when compared with the results of direct experiment on the paths of electron beams in the same field. The chief source of error is in determining, with templates, the radii of curvature of the equipotentials obtained by the electrolytic trough method. With less accuracy the same procedure can be used to trace rays farther removed from the axis, in order to investigate the spherical aberration of a lens (cf. section 39).

17. Ray-tracing by Other Methods: Automatic Tracing

A number of methods have been devised for finding electron trajectories from the differential equations of motion by introducing simplifying approximations. When the distribution can be analytically derived but direct integration turns out to be impossible, as in the case of the treatment of the two-tube lens (section 12 (*d*) above), then graphical methods can be used. Conversely, the axial distribution of the field, experimentally found, can sometimes be roughly represented by a function which allows of integration. A more general method of approach has been to assume the path of the electron in the plane of investigation to consist of arcs of circular, or parabolic, shape so that calculation of the steps in refraction are simplified. Alternatively, the curve for the axial distribution of potential, and its derivatives, can be broken up into short segments of simple shape so that a step-by-step solution of the differential equation of the ray is practicable. An example of this latter procedure will be given in the next section.

Salinger's method is typical of the former approach: it assumes the path of the electron to be comprised of arcs of circles of varying radius, the radius remaining constant between any two consecutive equipotentials assigned to the field. If in this region the component of the potential gradient normal to the path is E_n , then the local radius of curvature R must be given by

$$eE_n = mv^2/R;$$

and if V is the local potential with respect to the origin of the electron, we also have $eV = \frac{1}{2}mv^2$. Hence R is given by

$$R = 2V/E_n. \quad (\text{II.9})$$

A path segment of this radius is now drawn between the two equipotentials considered, being joined tangentially to the preceding segment, and the procedure is repeated from equipotential to equipotential through the whole field. This graphical method allows of rapid plotting of approximate trajectories, but is open to errors (which will be cumulative) in the junction of successive segments. For greater accuracy the whole process may be carried out numerically, but then becomes very laborious.

The chief interest of the method lies in its utilization as the basis of a machine for automatic tracing of trajectories, in conjunction with an electrolytic trough. In principle the procedure is to steer a trolley carrying the scribe X (Fig. 16) in such a way that its path is always of radius $R = 2V/E_n$, whilst it is moved forward in correspondence with the motion of the probe in the electrolytic tank. If this control can be smoothly maintained, the scribe will trace out the path of an electron in the given field.

As first described by Gabor, the scribe S was linked to the probe by a pantograph framework (Fig. 22). In order to measure the potential gradient E_n normal to the path, two auxiliary probes P_1 and P_2 are arranged equidistant on either side of the usual probe P ; they are kept aligned

parallel with the edge of the trolley T by the parallel link system shown. The central probe measures the local potential V , and the two outer probes measure the potential difference, and hence the gradient, normal to the electron path. The trolley is carried on two fixed wheels W_1 and W_2 on a common axis, and a third wheel W_3 on a movable axle. When in operation the wheels will be moving always parallel to the tangents to the instantaneous path, which has some point C as its centre of curvature. C may be found by producing the axis W_1W_2 to intersect the bisector of the angle between the plane of the wheel W_3 and the line joining S and W_3 . If the

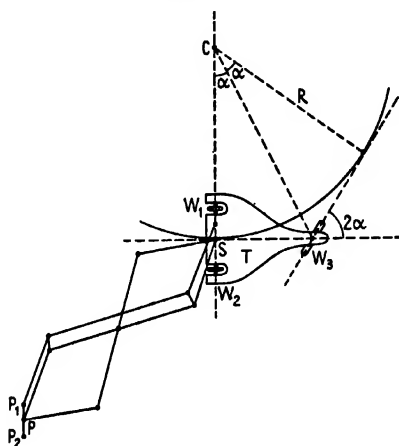


FIG. 22. Automatic ray-tracing machine (Gabor).

angle between the two axes of the trolley is then 2α , it follows that $R = SW_3 \cot \alpha$. If, therefore, the steering angle can be correctly related to the potential conditions, the path traced by S will have the radius of the trajectory; this will be so, from II.9, if

$$2V/E_n = SW_3 \cot \alpha. \quad (\text{II.10})$$

The necessary control of the steering is achieved by applying both V and E_n to a special bridge which shows a balance when II.10 is satisfied. The arm of this so-called 'tangent bridge' is connected by differential gearing to the wheel W_3 so that it is steered automatically as balance is maintained. The experimental procedure is to assume a direction for the incident electron and to start the probe moving into the field in the surface of the tank in this direction, the line P_1P_2 being kept normal to it; for instance, the path may initially be parallel to the axis. As soon as the field becomes measurable a normal gradient E_n will be detected, depending in amount on the curvature of the equipotentials. The bridge is adjusted to balance and the trolley and scribe correspondingly steered along the electron trajectory; and so on through the entire extent of the field.

Langmuir has devised a fully automatic version of this tracing machine

in which the trolley is driven by two synchronous motors mounted on it. One motor drives the fixed wheels, moving the trolley forward; the other drives a pin P that controls the steering angle of the third wheel (or bogie, as it is here) by varying the distance X (see Fig. 23). The motors are controlled by special potentiometer circuits actuated by the potential differences of the probes. Accuracy of the order of 1 per cent. is claimed, the chief error being in the determination of the gradient, E_n .

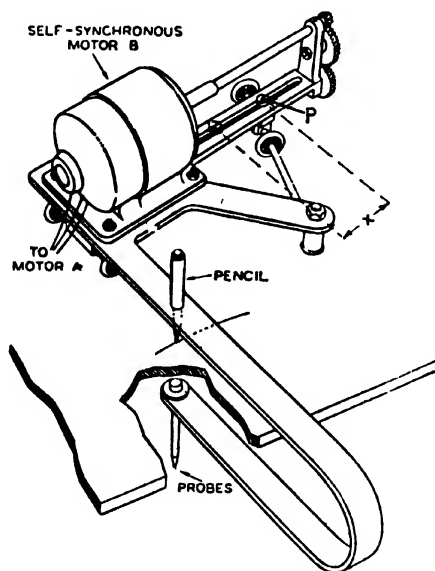


FIG. 23. Motor-driven ray tracer (Langmuir).

18. Deduction of Trajectories from Axial Potential Distribution

It will be understood from previous discussion that the focusing action of an electrostatic lens is determined by the potential gradient in the space between the axis and the bounding conducting surfaces; at any point this gradient will usually have components both parallel and normal to the axis. The shape of the bounding surfaces and the potential differences applied between them will determine the variation in the gradient from point to point, in much the same way as, in the analogous gravitational case, the rate of fall of a river and the general shape of its valley is related to the height and shape of mountain chains on each side. It should therefore be possible to calculate the field distribution around the axis from a knowledge of the symmetry conditions and the axial potential distribution; knowledge of the boundary conditions alone is, as a rule, insufficient for an exact solution, as already explained. Several attempts have been made to obtain trajectories by this method, but exact solutions are only possible when some tractable function can be fitted to the experimentally

determined axial distribution. In general, more or less drastic approximations have to be made according to the accuracy desired; usually the field, and therefore also the path of the electron, is considered as divided into a number of sections, as in the previous methods described. The variant given below has been in use in this laboratory for some time; it was developed by Dr. R. Sas Kulczycki, to whom the author is indebted for permission to include this unpublished work. It has the advantage of simplicity and speed in computation as compared with other methods.†

It is first necessary to obtain an expression for the field at any point off the axis in terms of the axial potential distribution and its derivatives. By neglecting higher derivatives than the second, and using approximate methods to find the first two, it is then possible to calculate the path through the lens of an electron entering in any given direction.

We start with an expression for the potential at any point (r, z) :

$$V_{(r,z)} = \frac{1}{2\pi} \int_0^{2\pi} f(z + ir \sin \alpha) d\alpha, \quad (\text{II.11})$$

which is a solution of Laplace's equation. That this is so may be shown by writing $z = (z_0 + a)$, expanding the function $f(z + ir \sin \alpha)$ as a power series, and finding the divergence of the gradient for the n th term:

$$\begin{aligned} \nabla^2(a + ir \sin \alpha)^n &= n(ir)^{n-2} \sin^{n-2} \alpha (n \cos^2 \alpha - 1) \\ &= n(ir)^{n-2} \frac{d}{d\alpha} (\cos \alpha \sin^{n-1} \alpha), \end{aligned}$$

a being set equal to zero.

Then integration gives $\frac{n(ir)^{n-2}}{2\pi} \int_0^{2\pi} d(\cos \alpha \sin^{n-1} \alpha) = 0$, and therefore

Laplace's condition is satisfied by the sum of all the terms, that is, by equation II.11.

We now expand the function $f(z + ir \sin \alpha)$ as a Taylor series:

$$f(z + ir \sin \alpha) = f(z) + ir \sin \alpha f'(z) + \frac{(ir \sin \alpha)^2}{2!} f''(z) + \dots + \frac{(ir \sin \alpha)^n}{n!} f^{(n)}(z),$$

which on integration term by term gives

$$\frac{1}{2\pi} \int_0^{2\pi} f(z + ir \sin \alpha) d\alpha = f(z) - \frac{r^2}{2!} f''(z) + \frac{r^4}{2! \cdot 4!} f^{(4)}(z) - \dots + \frac{(-1)^n}{(n!)^2} \left(\frac{r}{2}\right)^{2n} f^{(2n)}(z),$$

since all the odd terms are zero. As before, primes indicate derivatives with respect to z . We thus have the potential at any point (r, z) in terms of the axial distribution and its derivatives:

$$V_{(r,z)} = V_0 - \frac{r^2}{2!} V_0'' + \frac{r^4}{2! \cdot 4!} V_0^{(4)} - \dots \quad (\text{II.12})$$

† A similar method has recently been described by Goddard, in which, however, time-derivatives instead of z -derivatives are employed (cf. section 34.)

The radial electric field is at once found by differentiation with respect to r :

$$-\frac{\partial V}{\partial r} = E_r = \frac{r}{2} V''_{0z} - \frac{r^3}{2 \cdot 4} V'''_{0z} + \dots$$

When r is small, that is, when dealing with paraxial rays, only the first term need be considered for all practical purposes, so that

$$E_r = \frac{1}{2} r V''_{0z}. \quad (\text{II.13})$$

The radial force on the electron is then

$$m\ddot{r} = -eE_r = -\frac{1}{2} er V''_{0z},$$

and the radial acceleration

$$\ddot{r} = -\frac{1}{2} er V''_{0z}/m. \quad (\text{II.14})$$

But

$$\ddot{r} = d^2r/dt^2 = \frac{dz}{dt} \frac{d}{dz} \left(\frac{dr}{dz} \frac{dz}{dt} \right) = z \left(z \frac{d^2r}{dz^2} + \frac{dr}{dz} \frac{dz}{dz} \right),$$

whilst from energy considerations the axial velocity z is given by equation I.5:

$$z = v_z = \sqrt{\left(\frac{2eV_{0z}}{m} \right)};$$

whence

$$\ddot{r} = \frac{2e}{m} V_{0z} \frac{d^2r}{dz^2} + \frac{e}{m} \frac{dr}{dz} V'_{0z}. \quad (\text{II.15})$$

Combining II.14 and II.15 we get an equation containing only r and V_{0z} and their derivatives with respect to z :

$$4V_{0z} r'' + 2V'_{0z} r' + V''_{0z} r = 0, \quad (\text{II.16})$$

which is frequently referred to as the 'ray equation'. In theory the radial distance of the electron corresponding to any axial position z can then be found by integration:

$$r_s = - \iint \frac{r}{4V_{0z}} V''_{0z} dz dz - \iint \frac{1}{2V_{0z}} \frac{dr}{dz} V'_{0z} dz dz.$$

In practice, solution is attempted either by graphical integration and the use of methods of successive approximation, or by finding some integrable function that fits the axial potential distribution (usually of the form e^{-s^2/a^2}). We prefer to use an arithmetical step-by-step method of solution, after finding the axial potential distribution in the electrolytic trough; in the two-cylinder lens the distribution will be of the form shown in Fig. 19. The values of V_{0z} being thus known at intervals Δz along the axis (usually 1 cm.), its derivatives V' and V'' are then found from the first and second differences of this series of values, using expressions given by the calculus of finite differences; omitting the subscript $0z$ the complete expressions are:

$$V'_s = \frac{\Delta V_1 + \Delta V_2}{2\Delta z} - \frac{1^3 \Delta^2 V_2 + 3\Delta^2 V_3}{3! \cdot 2\Delta z} + \frac{1^3 \cdot 2^3 \Delta^3 V_3 + 5\Delta^3 V_4}{5! \cdot 2\Delta z} - \dots, \quad (\text{II.17})$$

$$V''_s = \frac{2 \Delta^2 V_1}{2! \Delta z^2} - \frac{2 \Delta^2 V_2}{4! \Delta z^2} + \frac{2 \cdot 2^3 \Delta^3 V_3}{6! \Delta z^2} - \dots, \quad (\text{II.18})$$

only the first term of each actually being employed. Here $\Delta V_1, \Delta V_2, \Delta V_3, \dots$ are the successive first differences in axial potentials ($V_2 - V_1, V_3 - V_2, V_4 - V_3, \dots$), and ${}^2\Delta V_1, {}^2\Delta V_2, \dots$ are the successive second differences ($\Delta V_2 - \Delta V_1, \Delta V_3 - \Delta V_2, \dots$ etc.), and so on for ${}^3\Delta V, {}^4\Delta V$, etc. The initial direction of an electron entering the lens being given, r_1 and r'_1 are known, and r'' can then be found from II.16. For finding the values r_2 and r'_2 at the end of the interval we have expressions corresponding to those for V' and V'' :

$$r'_2 = \frac{\Delta r_1 + \Delta r_2}{2\Delta z} = \frac{2\Delta r_1 + {}^2\Delta r_1}{2\Delta z}; \quad r''_1 = {}^2\Delta r_1 / (\Delta z)^2;$$

$$\text{giving} \quad r_2 = r_1 + r'_1 \Delta z + \frac{1}{2} r''_1 (\Delta z)^2, \quad (\text{II.19})$$

$$\text{and} \quad r'_2 = r'_1 + r''_1 \Delta z. \quad (\text{II.20})$$

These values of r_2 and r'_2 are then the commencing point for carrying out the same procedure in the next interval Δz . In this way the variation in r from segment to segment of the path can be calculated through the whole field; the greater the number of segments, that is, the smaller the interval Δz , the greater the accuracy of the result. When this is expressed in graphical form, a trajectory such as that shown in Fig. 31 is obtained.

For ease of working it is convenient to draw up the experimental data in tabular form and enter the values of V', V'', r, r' , and r'' as evaluated. An example is given in Table I, where values of the potential are given at intervals of 0.2 cm. along the axis of a two-cylinder lens of radius 1.0 cm. and length 3.0 cm. per cylinder, the voltages applied being 100 V. and 500 V. respectively. An electron is supposed to enter the low-voltage side parallel to the axis and at a distance of 0.08 cm. from it; that is, $r = 0.08$ cm. and $r' = 0.0$, as shown by the initial entries in columns 7 and 8. Then equation II.16 gives the value of r'' (column 9), and equations II.19 and II.20 give r and r' at the end of the segment, and so on. The values for r obtained in column 7 of the table are plotted against z in Fig. 31, and the trajectory produced to find the image point; this proves to be at $z = 4.98$ cm. The trigonometrical method of Klemperer and Wright gives the image point at $z = 6.2$ cm. The image actually found by Klemperer by direct experiment with such an arrangement was at 4.4 cm. The student will find it a valuable exercise to work out the path of an electron entering from the high-voltage side parallel to the axis and 0.08 cm. from it, as before, and then to plot both trajectories on the same graph. It will be found in this case that the image point occurs at 5.07 cm.; Klemperer found it at 5.2 cm. by experiment and at 5.7 cm. by the trigonometrical method of calculation. Experimentally the image distance for an accelerating ratio is always found to be smaller than for the corresponding decelerating case. The method of approximate calculation just described indicates that this should be so, whereas the method of Klemperer and Wright gives a result in the converse sense (6.2 and 5.7 in the present example). Mention will be made in a later section of the use of the two

methods in calculating the spherical aberration of such lens systems by tracing the paths of rays starting at various radial distances; the distribution of potential away from, as well as on, the axis must then be measured or calculated (cf. equation II.12).

TABLE I

Computation of Electron Trajectory in Two-Cylinder Lens for an Incident Beam parallel to the Axis at a distance of $0.08R$, where R is the radius of the cylinders; $V_1 = 100$ V., $V_2 = 500$ V., $\Delta z/R = 0.20$.

z/R	V (volts)	ΔV	ΔV	V'	V''	r/R	r'	r''
1.00	105.0	0.0800	0.0000	0.0000
1.10	105.3	0.3	0.6	3.0	15.0	0.0800	0.0000	-0.0029 _s
1.30	106.2	0.9	2.5	10.8	62.5	0.0799	-0.0008	-0.0123 _s
1.50	109.6	3.4	3.8	26.5	95.0	0.0796	-0.0031	-0.0177 _o
1.70	116.8	7.2	4.4	47.0	110.0	0.0786	-0.0066	-0.0180 _o
1.90	128.4	11.6	4.4	69.0	110.0	0.0769	-0.0102	-0.0143 _s
2.10	144.4	16.0	6.4	96.0	160.0	0.0746	-0.0132	-0.0177 _o
2.30	166.8	22.4	7.3	130.0	182.5	0.0716	-0.0167	-0.0134 ₇
2.50	196.5	29.7	7.3	167.0	182.5	0.0680	-0.0194	-0.0077 ₅
2.70	233.5	37.0	7.0	203.0	175.0	0.0640	-0.0210	-0.0029 ₇
2.90	277.5	44.0	1.0	223.0	25.0	0.0597	-0.0215	+0.0074 ₂
3.10	322.5	45.0	-1.0	223.0	-25.0	0.0556	-0.0201	0.0081 ₂
3.30	366.5	44.0	-5.5	206.0	-137.5	0.0517	-0.0184	0.0101 ₅
3.50	405.0	38.5	-6.0	178.0	-150.0	0.0482	-0.0164	0.0081 ₅
3.70	437.5	32.5	-12.0	133.0	-300.0	0.0451	-0.0148	0.0100 ₅
3.90	458.0	20.5	-6.9	85.3	-172.5	0.0423	-0.0128	0.0052 ₂
4.10	471.6	13.6	-3.8	59.5	-95.0	0.0399	-0.0117	0.0027 ₅
4.30	481.4	9.8	-2.9	41.8	-72.5	0.0376	-0.0112	0.0019 ₅
4.50	488.3	6.9	-2.6	28.0	-65.0	0.0354	-0.0108	0.0015 _o
4.70	492.6	4.3	-2.5	15.3	-62.5	0.0333	-0.0105	0.0012 ₅
4.90	494.4	1.8	-0.9	6.8	-22.5	0.0312	-0.0102	0.0004 ₅
5.10	495.3	0.9	0.0292	-0.0101	..

Then, mid-focal length $MF_1 = \left(\frac{292}{101} + 2.1\right)R = 4.98R$.

19. Direct Determination of Trajectories with a Gravitational Model

It has already been indicated that there is an analogy between the motion of a charged particle in a plane electrostatic field and that of a material body under gravitational forces. This analogy is exact, and the same methods can be used in both cases for determining field distributions and trajectories. Difference in height corresponds to potential difference, and mass to charge; the scale of the two systems is very different, but in principle there is complete correspondence. In a gravitational field (*in vacuo*) all particles of whatever mass have the same trajectory. It will be seen that equation II.16, for the path of the electron in an electrostatic field, does not contain either e , its charge, or m , its mass. As was to be expected, therefore, all charged particles will follow the same path in such fields whatever their value of e/m .

It follows that reference to a gravitational model of any electrostatic field will give direct insight into its deflecting properties; it is easy to visualize the path of a particle down a material slope, and also experimental tests of trajectories can be made directly. Working on any convenient geometrical scale from the electrode system, it is only necessary to ensure that the height of the model at any point is proportional to the potential of the corresponding point in the electrostatic field. The model may be built up from sheets of cardboard cut to the correct shape of the equipotentials, the thickness of a sheet corresponding to the potential interval. Alternatively, continuous slopes can be cast in plaster of Paris or a low-melting alloy. A photograph of a gravitational model of this type is shown in Plate I.

A much simpler and more adaptable procedure, when the field is two-dimensional, is to use a stretched rubber membrane to form the slope between points or surfaces at different gravitational potentials, as Kleyner first demonstrated. The paths taken by metal balls released from appropriate points can be observed or photographed, and the focusing properties of the system investigated directly. It is then an easy matter to vary the potentials, or even the shape, of the different electrodes and observe the effect on a trajectory.

The conditions which must be satisfied if the rubber model is to give a true reproduction of the motion of electrons in the parent electrostatic field have been thoroughly discussed by Zworykin and Rajchmann. They are: (a) that the slope of the surface must everywhere be small, and (b) that the tension of the rubber must be uniform over the whole surface. In addition, the radius of the rolling ball should be small compared with the radius of curvature of the surface, and its coefficient of friction with the surface should be as small as possible. It is convenient to reverse all potential differences, since the electron is a negative particle, and to make the height of any point in the model proportional to its negative potential.

Stout surgical rubber is used for the surface, about 1 mm. thick, stretched with tapes on a metal hoop, or pulled over and pinned to a square wooden frame (Plate II). Observation is assisted, and the uniformity of stretch readily tested, if squares are marked out on the surface in blue pencil before mounting it. When recording the trajectories photographically, contrast is improved by using black rubber. The electrode models are shaped from lead, or light-gauge aluminium, to the form of a longitudinal section of the lens system. The model of the 'lens' of Fig. 12 will therefore consist of two identical pairs of parallel plates, but necessarily of finite length; for accurate reproduction of the field the length should be large compared with their separation. Each pair is mounted firmly on a base plate and raised beneath the rubber surface to a height proportional to the negative potential applied to the original electrode, according to the vertical scale adopted in the investigation. The longitudinal gap between the two pairs will likewise be equal to the separation of the two sets of plates in the original. More usually the dimensions of the model will be

several times those of the electrode system, to scale, in order to allow of more detailed investigation; the rubber model can readily be made 2-3 ft. each way.

For convenience of working, the models of the electrodes are first disposed at the appropriate heights on a level surface, such as a stout sheet of glass, and the rubber surface then pressed down upon them. To ensure close contact of the rubber with the edges of the model at all points it is frequently necessary to employ counterparts pressing down from above, especially on the more positive lower electrodes, or suction tubes. The surface then takes the shape corresponding to the potential distribution between the real electrodes, the slope at any point being proportional to the field strength, and the force exerted by an electrode on the rubber

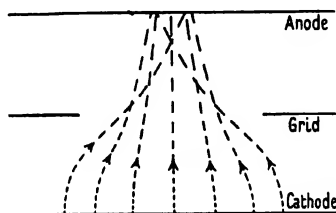
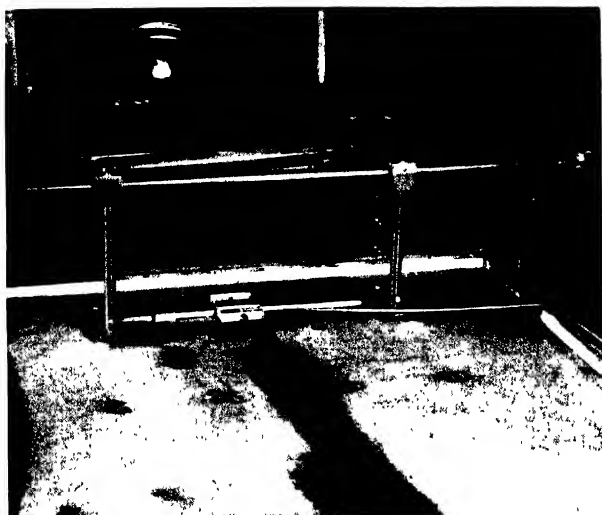


FIG. 24. Trajectories of ball-bearings (\equiv electrons) in rubber model of plane 3-electrode system (after Kleynen). (Original is photograph obtained with intermittent illumination.)

being proportional to the capacity of that electrode. If a steel ball is now released at the highest end of the system, or projected into it from a chute, on or near the axis, its trajectory over the surface can be observed; care has to be exercised that it is not given an unintentional initial velocity transverse to the axis, as this would produce considerable error in the path.

It is often sufficient to note only the point at which the particle crosses the axis in the lower half of the 'lens', and to mark this 'image point' on the surface for several initial directions of projection; the approximate focal length and total aberration of the system is thus obtained. If a study of the whole path is required, the surface is illuminated from above with lamps lit from the 50-cycle mains. When a camera is exposed over the surface during the motion of the steel ball across it, its path will be recorded as a series of dashes of light produced by the reflection from it of the intermittent illumination. The separation of these dashes will increase with the speed of the ball, so that the photograph gives useful information as to its velocity in different parts of the lens, as well as the complete path (Fig. 24). Steel ball-bearings of 4 or 5 mm. diameter are of suitable size for this purpose.

The accuracy of such a gravitational investigation is naturally less than that of the best numerical methods of ray-tracing, but it gives the main properties of any system directly and with little labour. Its main advantage is the speed with which the variation of potential and electrode form



a. The ball-release mechanism (--- cathode), and the three rows of grid 'wires'.



b. General view, showing arrangements for photography of tracks. The 'anode' bar is also visible.

PLATE II. RUBBER MODEL OF POTENTIAL FIELD IN A (PLANAR) PENTODE.
(The Mullard Radio Valve Co. Ltd.)

in complicated systems can be investigated. As an instance, Zworykin and Rajchmann used it to find the form of electrode required in a multi-stage secondary emission electron multiplier in order to bring all the electrons from each surface on to the next, under the influence of electrostatic fields alone. The final arrangement is described later, and shown in Fig. 100; it will be appreciated that the mathematical solution of such a problem would have been impossible, even by approximate methods.

The rubber model, therefore, offers a most flexible and rapid means of investigating electron trajectories. For many purposes it will give sufficiently accurate results even if no great care is taken to ensure uniform stretching of the surface, and if the slopes are made as steep as 30° to 40° , in order to increase the speed of the steel ball.

FURTHER READING

1. Solution of Electrostatic Problems:

Principles of Electricity (Page and Adams), Chapter III.
Christopherson and Southwell, *P.R.S.* **168**, 317, 1938.
Bertram, *P.I.R.E.* **28**, 418, 1940; *J. App. Phys.* **13**, 496, 1942.
Motz and Klanfer, *Proc. Phys. Soc.* **58**, 30, 1946.

2. Electrolytic Trough and Automatic Ray-tracing:

Gabor, *Nature*, **139**, 373, 1937.
Langmuir, *ibid.*, p. 1067, 1937.

3. Ray-tracing Methods:

Klemperer and Wright, *Proc. Phys. Soc.* **51**, 296, 1939.
Spangenberg and Field, *P.I.R.E.* **30**, 138, 1942.
Goddard, *Proc. Phys. Soc.* **56**, 372, 1944.

4. Rubber Model:

Kleynen, *Philips Techn. Rev.* **2**, 338, 1937.
Zworykin and Rajchmann, *P.I.R.E.* **27**, 558, 1939.

CHAPTER III

ELECTROSTATIC FOCUSING

20. *Types of Lenses*

IN principle, any arrangement of electrodes will function as an electron lens if they produce regular curvature of the equipotential surfaces. The arrangement of two identical coaxial cylinders separated by a small gap, which has been discussed above, is the simplest type of focusing system. Many variations of the two-cylinder form are possible, and other types have also been devised. They differ primarily in the general form of the equipotential map, which is essentially determined by the geometry of the electrodes. They are therefore usefully classified as follows:

(a) *Two-cylinder Lenses*. Some of the possible forms are shown in Fig. 25. The cylinders may be of different diameters, may be fitted with diaphragms pierced by apertures of various sizes, and may interpenetrate each other. The equipotentials are symmetrical in appearance only, and not quantitatively. It is this type which is most used in the focusing systems of cathode-ray oscilloscopes. The optical analogue is the meniscus lens.

(b) *Aperture Lenses* are distinguished by the existence of uniform fields on either side of an aperture, within which a transitional region is produced. When the electron approaches from the side of lower field strength the action is converging; and, vice versa, it is diverging when the electron moves in the opposite sense. The latter arrangement is the only system that, over all, has a diverging effect on an electron beam; its power, however, is very limited. These forms, and the shape of the equipotentials near the aperture, are shown in Fig. 37; they are sometimes referred to as 'pinhole' lenses, and are usually employed in conjunction with one of the other types of lens.

(c) *Symmetrical, or 'Univoltage', Lenses*. This type is a development of the two-cylinder system, a third cylinder being introduced between two identical cylinders, which may be longer or shorter than the middle one (Fig. 26). As shown, the cylinders are sometimes reduced to simple apertures. The central element may be at higher or lower potential than the outer two, which are maintained at the same potential. If this common potential is zero (earth) and the central element is connected to the cathode of the system, then only a single (negative) voltage need be applied, and the arrangement is known as a 'univoltage' or 'einzel' lens; or, of course, the cathode and central element may be earthed, and the other two electrodes kept at high positive potential. The equipotentials are entirely symmetrical about the mid-plane of the system, and hence it is sometimes also known as the 'symmetrical' lens; the refractive indices on the two sides of the lens are identical. The optical equivalent is a

combination of two concave with a convex lens, or vice versa, depending on whether the central element is at lower or higher potential than the other two; in either case the total effect is convergent. Its most important application is in the electron microscope.

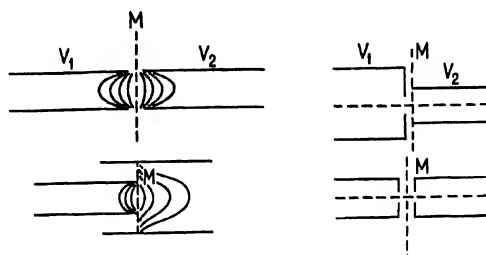


FIG. 25. Types of two-cylinder lens.

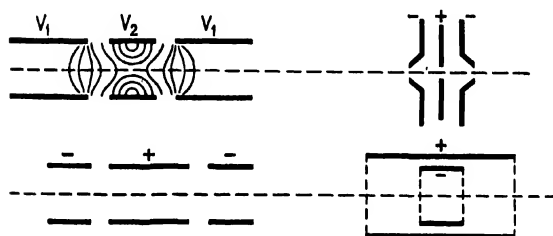


FIG. 26. Types of symmetrical lens.

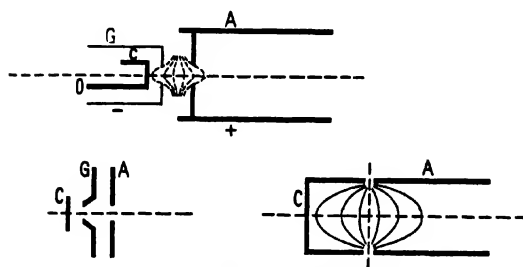


FIG. 27. Types of immersion lens.

(d) *Immersion Lenses.* Properly speaking, an immersion lens is formed by any system in which object and image are in regions of different refractive index, and in this sense both the two-cylinder lens and the aperture lens are of the immersion type. In general usage, however, the term is applied to arrangements of three electrodes at different potentials, one of which is

the source of electrons, i.e. the cathode. One of the electrodes must be at higher potential (anode, *A*), and the other at intermediate or lower potential (grid, *G*), than the cathode *C*. The 'object' is thus well immersed in the focusing field of the 'electron gun', as it is often called (Fig. 27). Optically it is equivalent to a meniscus lens; but its power may be so great that the rays cross over within the lens field itself. It constitutes the usual source of electrons in the cathode-ray tube and many other devices. In a special form the grid may be joined to the cathode, giving effectively a two-cylinder system, as with the early Wehnelt cylinder and in Zworykin's image converter.

21. Cardinal Points

In considering the imaging of objects by a lens of whatever type, it is inconvenient to have to trace the ray-paths from a particular position of the object in order to find the resulting position and size of the image. As in light optics, a set of six points may be defined from a knowledge of which it is immediately possible to calculate the solution to such a problem. These are the cardinal points of the lens system, being the two focal points, the two principal points, and the two nodal points. When they have been found, either by tracing representative rays by the methods given in Chapter II or by direct experiment, the position and size of the image may be deduced for any position and size of the object, no matter what the number of the components of the lens system or the complexity of the refracting field. The procedure, however, is only exact for paraxial rays: that is, for conditions in which the formative rays are inclined at a small angle θ to the axis. In such circumstances the higher terms in the expression for the sine of the angle

$$\sin \theta = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots \quad (\text{III.1})$$

may be neglected and the angle equated to its sine. A powerful method is then available for describing the characteristics of lens systems. In this approximate form the treatment is usually called Gaussian optics, after its originator, or alternatively the 'first-order theory', in contradistinction to a more rigorous 'third-order theory' which attempts to include the second term of the sine series. The nature of the cardinal points will be briefly recapitulated here and the formulae expressing their interrelationship stated; the derivation of the latter will be found in any standard work on optics. The sign convention used will be that of the normal Cartesian system: distances measured to the right from any reference point or plane will be positive and to the left negative. As in glass optical practice, a converging lens will be positive in power and a diverging lens negative.

Suppose some lens system to be located on the axis *z* and to possess rotational symmetry about it (Fig. 28); it is usual to specify its position

by its geometrical mid-plane M , or some such convenient plane of reference. The exact form of the lens or of its components is immaterial, but for simplicity it may be assumed to be convergent. Now suppose a ray AB , parallel to the axis, to enter the lens from the left. It will be refracted and emerge in some direction F_2X , cutting the axis at F_2 . On being produced backwards it will intersect its initial direction in some point P' . Then the plane through P' normal to the axis is the principal plane, and its point of intersection (P_2) with the axis is the principal point, of the image space; by convention they are styled the *second* principal plane and point. For all rays entering from the same direction the refracting influence of the lens may be supposed to be concentrated in this plane. In the first-

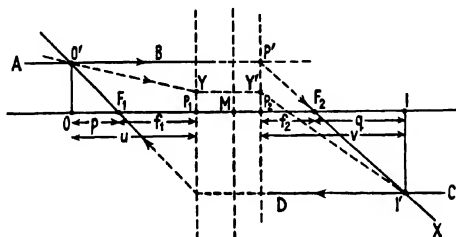


FIG. 28. Principal points of a lens.

order theory, or, in other words, in the absence of aberration, all rays parallel to AB will cut the axis in the point F_2 after refraction. This is the focal point, and the plane through it normal to the axis is the focal plane, of the image space. Alternatively we may say that the image will lie in this plane when the object is infinitely distant from the lens.

By considering a second ray CD entering parallel to the axis from the other side of the lens we define the focal point (F_1) and focal plane, and the principal point (P_1) and plane of the object space. Let these two rays intersect on the one side of the lens at O' and on the other side at I' ; drop perpendiculars $O'O$ and $I'I$ respectively on to the axis. Then $I'I$ can be considered to be the image of an extended object situated at OO' . The graphical construction of an image of a given object, given the position of the principal planes and focal points, should now be clear. The ray from O' parallel to the axis will leave the second principal plane at the same distance from the axis as that at which it intersects the first, passing then through the second focal point. The ray from O' through the first focal point will proceed parallel to the axis from the point at which it intersects the first principal plane. The intersection of these two rays gives the image point I' corresponding to O' . Any other ray $O'Y$ will intersect the principal planes according to the same general rule: it will cut the second principal plane at the same distance from the axis as that at which it met the first principal plane. In other words, these are planes of unit lateral magnification, and this constitutes their simplest definition.

The distance F_1P_1 is known as the first focal length and F_2P_2 as the

second focal length. Then if $F_1O = p$ and $F_2I = q$, we have for the lateral magnification

$$M = \frac{y_2}{y_1} = \frac{II'}{OO'} = \frac{f_1}{p} = \frac{q}{f_2},$$

whence

$$f_1 f_2 = pq, \quad (\text{III.2})$$

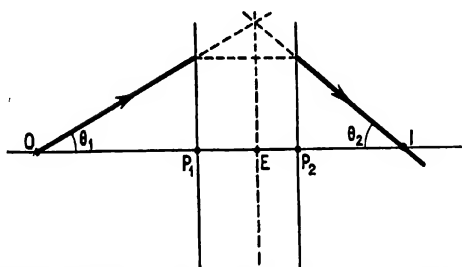


FIG. 29. Angular magnification and equivalent thin lens of a lens system.

as was originally shown by Newton. Measuring the object and image distances, u and v , from the principal planes we also have

$$\frac{f_1}{u} + \frac{f_2}{v} = 1. \quad (\text{III.3})$$

Either of these relations allows us to calculate the position of the image, in preference to its graphical location. It is also of value to define a quantity named the angular magnification, in terms of the angles θ_1 and θ_2 made with the axis at the object and image respectively by any ray passing from O to I (see Fig. 29):

$$\frac{\theta_2}{\theta_1} = \frac{u}{v} = \frac{p}{f_2} = \frac{f_1}{q}. \quad (\text{III.4})$$

It may be shown that the angular and the lateral magnification are inter-related, as expressed in Lagrange's law:

$$N_1 \theta_1 y_1 = N_2 \theta_2 y_2, \quad (\text{III.5})$$

where N_1 is the refractive index of the object space and N_2 that of the image space; in electron optics they are usually not the same, being proportional to the square root of the respective potentials (V_1, V_2) in the two spaces. It follows, by combining the above expressions, that

$$\frac{f_1}{f_2} = \frac{N_1}{N_2} = \sqrt{\frac{V_1}{V_2}}; \quad (\text{III.6})$$

so that in a lens of given voltage ratio the measurement of one focal length allows of the calculation of the other (cf. p. 66). It follows also, from III.5 and III.6, that the lateral magnification is given by

$$M = \frac{y_2}{y_1} = \frac{v N_1}{u N_2} = \frac{v}{u} \sqrt{\frac{V_1}{V_2}},$$

and depends on the root of the voltage ratio applied to the electrodes.

For some purposes it is convenient to think in terms of the equivalent thin lens which would have the same refracting effect as the 'thick' lens under consideration. The position of this equivalent lens is given by the plane E (Fig. 29), drawn normal to the axis through the point of intersection of the directions on either side of the lens of any ray passing from the axis at O to the axis at I . Obviously the position of this lens will vary with the position of the object and image, but it always lies between the two principal planes. From Fig. 29 it will be seen that

$$\frac{EO}{EI} = \frac{OP_1}{IP_2} = \frac{u}{v}, \quad (\text{III.7})$$

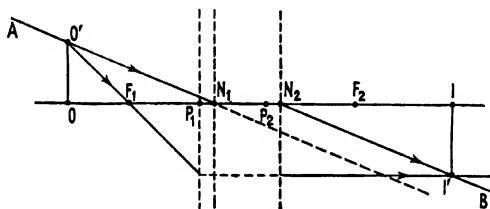


FIG. 30. Nodal points of a lens.

from which the position of the equivalent thin lens may be calculated corresponding to any position of the object.

In many practical applications it is necessary to have a ready indication of the position of the focal points with respect to the geometry of the electrode system. For this purpose the 'mid-focal length' is utilized, being the distance from the focal point to the mid-plane of the lens (M , Fig. 28). When the distance MF is quoted, it is usually expressed as a multiple of the radius of the lens cylinder in which it falls; it should also be stated whether it is obtained with the electron beam accelerated or decelerated by the potential ratio applied to the lens.

It remains to define the nodal points and planes. As the principal planes are planes of unit lateral magnification, so the nodal planes are of unit angular magnification. An incident ray that is directed towards the first nodal point on the axis (N_1 in Fig. 30) will leave the lens in a direction passing through the other (N_2), and parallel to the incident ray; the ray AN_1 will emerge in the parallel direction N_2B . It will be seen that these two points now share the property of the single 'optical centre' of a thin lens. The planes normal to the axis through N_1 and N_2 are the nodal planes. They may be used, instead of one of the focal points, in constructing the image as shown. With respect to the position of the corresponding focal points, they lie in the same direction as the principal points but the separations are interchanged:

$$N_1F_1 = P_2F_2 \quad \text{and} \quad N_2F_2 = P_1F_1. \quad (\text{III.8})$$

Therefore the nodal points may be located as soon as the other four cardinal points have been found. Further, just as the angular magnification

is related to the distance of object and image from their respective principal points, it can be seen that the lateral magnification is related to their distance from the respective nodal points:

$$M = \frac{y_2}{y_1} = \frac{II'}{OO'} = \frac{IN_2}{ON_1}. \quad (\text{III.9})$$

Finally it may be mentioned that when the indices of refraction of the object and image space are the same, as in most light-optical problems and in the case of the univoltage lens of electron optics, we have $f_1 = -f_2 = f$, and $1/v - 1/u = 1/f$, as in a thin lens; also the nodal and principal points coincide.

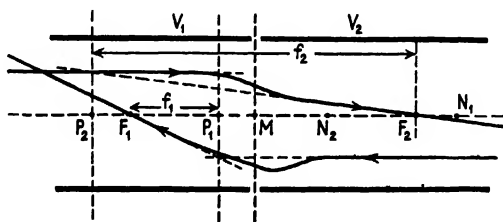


Fig. 31. Trajectories and cardinal points in a two-cylinder lens ($V_2 > V_1$).

22. Cardinal Points of Electron Lenses

The position of the cardinal points for any lens, whether in light or electronics, may be found by ray-tracing methods. It is only necessary to trace the path through the system of two representative rays, incident parallel to the axis from opposite directions. These serve to locate the two focal points and the two principal points; the nodal points may then be located by the use of expression III.8. Once these have been found it is possible to draw directly, or to calculate, the paths of any other rays through the lens; or, rather, to find the emergent direction for any incident ray, as the detailed path in the lens is not determinable.

If we draw in this way the representative trajectories for the simple two-cylinder lens we find the positions of the cardinal points to be as shown in Fig. 31. The principal and the nodal points are crossed over with respect to the corresponding focal points. This situation follows from the 'meniscus' character of the lens: the ray entering from the low potential side suffers deviation away from the axis in the second half of the lens, and, therefore, must be produced backwards correspondingly farther before it intersects its initial direction, to give the position of the second principal plane. Consequently this plane falls in the object space, and not in the image space, as is usual. On the other hand, a ray approaching from the high potential side is first deviated from the axis, and then towards it in the later part of its path; it soon cuts its initial direction when produced backwards, and the first principal plane falls in the object space, in the

usual way, but not so near the object as the second plane. It follows that the nodal planes are similarly crossed over in the image space. This condition is found in all two-cylinder lenses: the principal planes are crossed over on the low-voltage side of the lens, and the nodal points are crossed over on the high-voltage side. The expressions III.2 to III.9 given above are unaffected, of course, and may be used as indicated to locate the image.

In discussing the focusing power of such a lens it is useful to consider it as composed of two semi-lenses of opposite power. When the voltage ratio is greater than unity, the first cylinder has a converging and the second a diverging action on the electron beam, the optical analogue being a meniscus lens. At first sight it might be thought that the two semi-lenses would neutralize each other, since the equipotential plot is symmetrical, and it is important to understand why this is not so. The power of any lens component depends on the voltage ratio across it, and the potential in the mid-plane is $V_M = \frac{1}{2}(V_1 + V_2)$, if the overall voltage ratio of the system is V_2/V_1 . Hence the power of the positive semi-lens depends on the ratio

$$\frac{V_M}{V_1} = \frac{V_1 + V_2}{2V_1},$$

and that of the succeeding negative semi-lens on the ratio

$$\frac{V_2}{V_M} = \frac{2V_2}{V_1 + V_2}.$$

Clearly, as the voltage ratio is increased, the first expression increases without limit, whereas the second approaches the value 2. Thus the positive component is always stronger than the negative, and the power of the whole system increases with the voltage ratio. When the ratio has passed the value 10:1, the second component has little further influence on the focusing power of the lens as a whole.

23. *Experimental Location of Cardinal Points*

However accurate ray-tracing methods may be, it is essential to check the information which they yield by direct investigation of the paths of electron beams through the given lens system. Experimental work of this nature is made difficult by the necessity of maintaining the whole of the electron path in high vacuum. Attempts have been made to follow the trajectories by using such a pressure in the tube (of the order of 10^{-3} to 10^{-2} mm. mercury) that they are rendered visible by the luminous excitation of gas molecules. Whilst such methods are of value for qualitative examination of electron paths in various devices, they are of doubtful validity when it is a question of the location of cardinal points.

Experimental arrangements of some accuracy have been devised, however, by Epstein and by Klemperer and Wright, among others. The apparatus of the latter workers will be described here. It consists essen-

tially of an evacuated glass tube G (Fig. 32 *a*), which contains the electrode system to be investigated, an electron gun to project parallel beams into it, and a 'target' or viewing screen. In order to protect the beams from disturbance by external magnetic fields (including that of the earth) the apparatus is surrounded, when in use, by an outer tube of a material of high permeability, such as Mumetal. The electron gun E is specially designed to give a parallel beam of large diameter from the concave surface of a hot cathode C . Smaller pencils of rays are then selected by means of a 'pepper-pot' diaphragm D , shown to a larger scale in Fig. 32 *b*. The sets of holes situated at different distances from the axis may be uncovered as desired by manipulating the flap B , connected to a rotatable ground-glass joint. Parallel pencils of paraxial, marginal, or intermediate character can be produced by this means for investigating the aberration of the lens. They pass through the focusing cylinders L_1 and L_2 and impinge on the target T , which consists of a glass plate coated with a fluorescent screen that can be moved to and fro along the axis by magnetic control. When accurate measurements are needed a graticule of wire gauze is embedded in the glass plate of the target. This is viewed from the side, through a right-angled prism fixed to the glass end-plate that seals the tube, by means of a microscope M having a scale in its eyepiece.

The procedure with each set of rays, and with a given potential ratio on the electrodes, is to move the target until the sharpest focal spot F is observed in the eyepiece. By careful measurement of the size of the spot as the target is moved towards the lens it is then possible to determine the emergent direction of the rays. These readings are plotted to scale on a longitudinal section of the lens system, and by producing the rays backwards through the lens the points of intersection with the incident pencils (i.e. the principal planes) are located. The second focal point and principal plane being thus determined, the voltage ratio can then be reversed and the corresponding points for the image space found, supposing the geometry of L_1 and L_2 to be symmetrical; otherwise the whole lens system would have to be reversed. The position of the nodal points can then be calculated from equation III.8.

A number of precautions must be observed if reliable results are to be obtained. In the first place, the gap between the electron gun and the lens must be electrostatically shielded, a cage being fitted around it as shown in the diagram (S). Similarly it is necessary to shield the gap between the two elements of the lens, L_1 and L_2 , and this is achieved by fitting a 'skirt' to the first cylinder. This gap should in any case be small, and preferably less than one-tenth the radius of the cylinders, otherwise the potential field within the lens departs appreciably from that assumed in ray-tracing methods. For the same reason the cylinders must be sufficiently long for the focusing field not to be influenced by either the pepper-pot diaphragm or the target. If the length is not appreciably greater than the radius, the outer equipotentials are distorted, and these are most important in determining the aberration of the system. It is found that the inter-

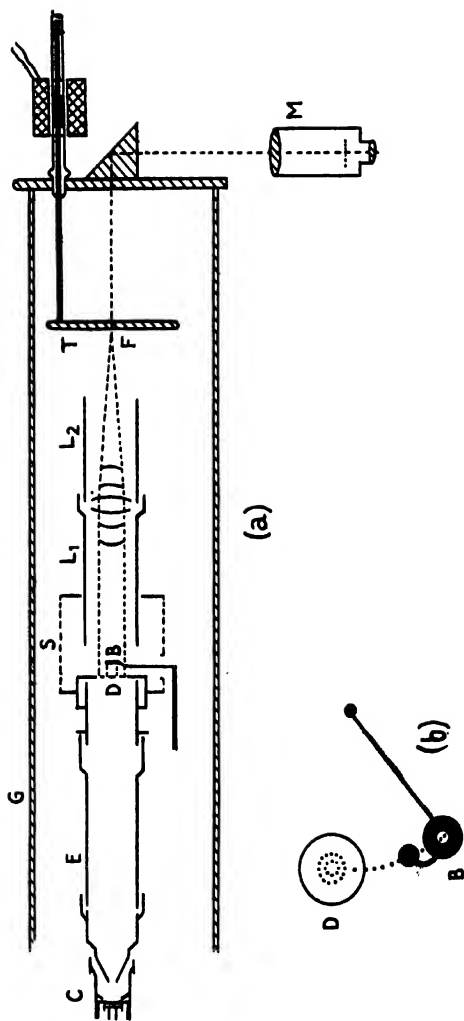


Fig. 32. Experimental investigation of electron trajectories (Klemperer).

ference is negligible if the length of each cylinder is made at least three times the radius. The chief source of error lies in the great depth of focus of paraxial rays, under an accelerating potential; it is obvious from Fig. 31 that the emergent ray in this case crosses the axis at a very small angle. Consequently the diameter of the spot on the target varies but little over an appreciable axial distance, and the position of the focus is difficult to determine exactly. For rays entering the lens at a greater radial distance, and for all rays in a decelerating system, the agreement between experiment and ray-tracing is better. In strong lenses, with a high voltage ratio, the focal point may lie within one of the lens cylinders, which must be kept long for the reason given above. In such a case no attempt may be made to follow the beam back into the cylinder with the target, as the field would be greatly disturbed in the process. The procedure used is to take accurate readings of the divergence of the beam beyond L_2 , and then to plot the values to scale and extrapolate to minimum diameter of beam.

Details of the results obtained with this apparatus will be given in the next section, and later in connexion with lens aberrations. It is usually sufficient to record the variation of the mid-focal distance only (MF), and not of all the cardinal points, for different rays and focusing ratios.

24. *The Two-cylinder Lens*

This type of lens was first devised and investigated by Zworykin and his collaborators. The nature of its field and of representative trajectories in it have already been indicated. It remains to discuss the focusing characteristics of the system and their dependence on the geometrical and electrical variables.

A number of attempts have been made to derive theoretically an expression for the focal length in terms of the voltage on one or other of the electrodes and the potential variation along the axis. Unfortunately a simple result can only be obtained by applying approximations that are valid only for voltage ratios little greater than one; the full expression, on the other hand, is unwieldy (see section 26).

For a given pair of electrodes, of the form shown in Fig. 25, the focal length is found to depend on the voltage ratio applied to them and on their radii. If the actual voltages are varied, but the ratio kept the same, the focus also remains the same. If the ratio is kept constant for systems of different radius, the focal length changes in proportion to the radius. The cardinal positions and the mid-focal distance, therefore, are usually given in terms of the tube radius; so that, for instance, a second focal length of $2.5R$ would correspond to an actual distance of 5 cm. in a lens of 4 cm. diameter and to 10 cm. in one of twice the size.

The essential characteristic of any lens is, therefore, the relation between the voltage ratio and its focal length or mid-focal distance. Fig. 33 shows the form of curve obtained as the ratio is varied, both when the electrons are being accelerated (A) and decelerated (D). The mid-focal distance on the high-voltage side is found always to be less than on

the low side, and the difference increases with the voltage ratio. As Fig. 31 shows, however, the second focal length proper (f_2) will be much larger than the first (f_1), even though the mid-focal distances are nearly equal, because the principal planes are crossed over. In Fig. 34 this is shown by plotting f_1 and f_2 against the voltage ratio. All these curves show very little further decrease in focal length above a ratio of 5 or 6. This property is even more evident in the case of the principal planes (Fig. 35) in both

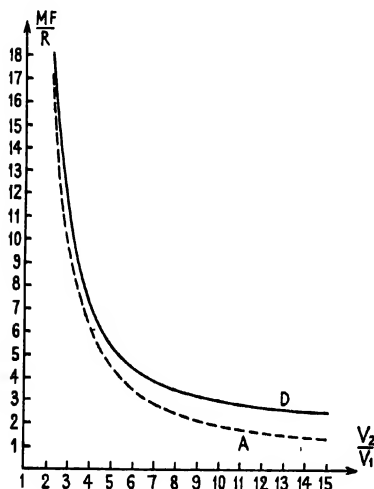


FIG. 33. Mid-focal length (MF) of two-cylinder lens, with varying voltage ratio (Klemperer).

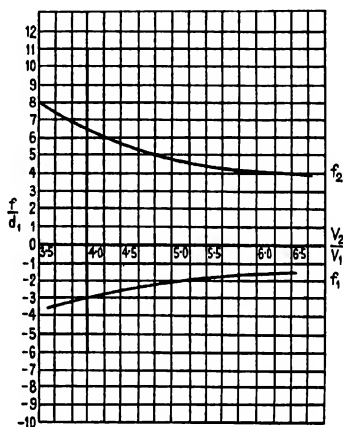


FIG. 34. Focal-length (f_2 and f_1) of two-cylinder lens, with varying voltage ratio (after Epstein).

(d_1 = diameter of first cylinder.)

the accelerating and decelerating cases. The lens cylinders in these experiments having a length of at least three times their radius, it follows from the curves that both the first focal point and the second principal plane fall well back in the low-voltage cylinder or even beyond it. On the high-voltage side the focal point lies outside the lens for low voltage ratios, but retreats into the tube for ratios above 5 or 6. When the ratio is very high, the rays may cross over so close to the mid-plane that the beam leaves the lens more divergent than on entering; if it reaches values of the order of 1,000, as in some electron guns, the beam may in fact cross the axis twice within the lens.

The effect of increasing the diameter (d_2) of the second cylinder, the diameter (d_1) of the first being kept constant, is much the same as that shown on decreasing the voltage ratio. Typical curves of the focal length against the ratio of the diameters, for a constant voltage ratio of 1:4, are reproduced in Fig. 36. In experiments with this type of lens, where there is no mid-plane of symmetry, the mouth of the smaller cylinder is taken as the plane of reference for measuring the 'mid-focal' distance, MF .

Clearly the properties of the different possible variants of the symmetrical two-cylinder lens are not essentially different from those of the parent. However, if we consider the lens as composed of a positive and negative semi-lens, it is surprising that the change in focal length is as great as indicated by these curves. As was shown in section 22 the first, low-voltage,

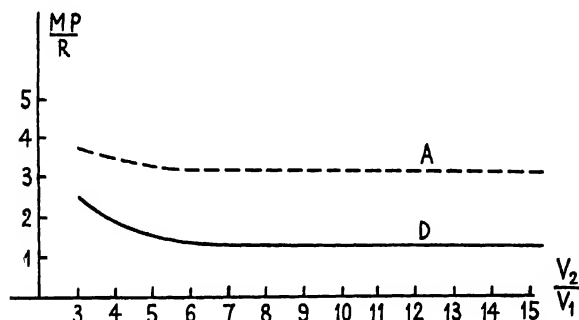


FIG. 35. Variation of position of principal planes with voltage ratio in two-cylinder lens (Klemperer).

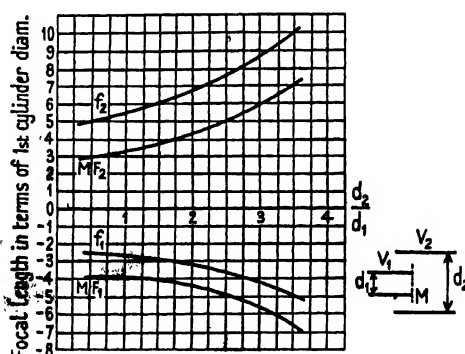


FIG. 36. Variation in focal length with ratio of tube diameters (d_2/d_1) in two-cylinder lens; $V_2/V_1 = 4$.

semi-lens increases in power without limit as the voltage ratio rises, whilst the second semi-lens, on the high-voltage side, tends to a limiting ratio of 2. The focusing action of the lens thus depends largely on the first semi-lens, on the assumption that the 50 per cent. equipotential lies in the mid-plane. It might be expected, therefore, that increasing the diameter of the second cylinder would have little effect on the power of the system; if anything, a small increase in the second focal length would be expected, on account of the general rule of increase in focal length with diameter in the symmetrical lens. The curves show, however, that not only does f_2 increase appreciably (from 5.5 to 10, as d_2/d_1 rises from 1 to 3.5), but f_1 suffers an increase (negatively) that proportionately is as great (from

—2.7 to —4.9). The divergent semi-lens has increased, and the convergent decreased, in power on increasing the diameter of the former. Klemperer has pointed out that this is due to the equipotentials bulging farther into the larger tube in such a case. The 50 per cent. surface, which lay close to the end of the low-voltage tube in the symmetrical arrangement, now lies well inside the other tube. On doubling the voltage ratio the 31 per cent. equipotential is found at M , and at four times the ratio the 13 per cent. surface lies there. The positive power in the smaller tube is decreased, and the negative power of the larger increased, by this change in the voltage ratio in each semi-lens. Conversely, on decreasing the diameter of the low-voltage cylinder the focal length is not decreased in exact proportion, for the same reason; but the change is considerably greater than that produced by a similar decrease in the diameter of the high-voltage cylinder alone. Detailed information on the focal curves obtained with various relative dimensions and dispositions of the electrode system will be found in Maloff and Epstein's book, *Electron Optics in Television*, and will not be further discussed here.

The lateral magnification M produced by such lenses is given by

$$M = \frac{v}{u} \sqrt{\left(\frac{V_1}{V_2}\right)}.$$

Since the principal planes are crossed over well back in the object space, v is considerably greater than the image distance measured to the mid-plane, and u is much less than the object distance from this plane. However, the inverse of the root of the voltage ratio enters into the expression for M , and the result is that the magnification is not very different from the simple ratio of the image and object distances measured from the mid-plane: $MI/MO = [(MF_2 + q)/(MF_1 + p)]$. Klemperer states that the measured magnification is usually about 70 per cent. of this ratio; Spangenberg and Field give a figure varying from 60 to 82 per cent. for different forms of the two-cylinder lens, and as high as 95 per cent. for some types of aperture lens.

Lenses composed of more than two cylinders have much the same properties as the simple parent type, so long as the potential is applied in successive steps from each tube to the next. The focal length of a chain of n tubes, however, is found to be longer than that of a two-cylinder lens to which the same total voltage ratio, $V_n:V_1$, is applied. As the number of electrodes is increased, the focal length continues to increase indefinitely; in the limit, with an infinite number of tubes, there will be acceleration without focusing of the beam. This condition may be reached by coating the interior of a glass tube with a thin layer of metal or of graphite, the high resistance of which will produce a uniform potential gradient between contact electrodes applied at either end. The system is then no longer a lens but an accelerator tube. On the other hand, if in a multi-cylinder lens one of the intermediate tubes is set at a lower potential than both its immediate neighbours, a different equipotential distribution is formed

from that described above, and the focusing action differs correspondingly. The lens is then similar in character to the 'symmetrical' or saddle-field lenses discussed in section 26.

25. Aperture Lenses

An aperture lens is formed whenever a diaphragm, pierced by a small hole, is inserted into the previously uniform field existing between two electrodes at different potentials. The diaphragm may be at the same potential as one of the electrodes, or at some intermediate potential. If it had no aperture, it would have the effect of producing a uniform field on either side, and the equipotentials would still be straight; in general the two parts of the field would now have different gradients, so that the equipotentials would be differently spaced on the two sides of the diaphragm. The presence of the aperture results in the stronger field exerting an influence over a small region on the farther side, into which its equipotentials now protrude. The two most usual arrangements are shown in Fig. 37 (*a* and *b*), together with the variation of the potential and of its first derivative ($\partial V/\partial z = -E$) along the axis. The characteristic feature is that the field rises sharply over a small distance about the aperture. Frequently in case (*a*) V_1 is made equal to V_2 , and in case (*b*) V_1 is made equal to 0; the essential action of each lens is not thereby altered. It is possible, of course, for the central or final electrode to be negative, as indicated in Fig. 37 (*c* and *d*). These arrangements would be ineffective for electrons emitted from the first electrode, owing to the large decelerating field, but they can be used when the beam has been previously accelerated and enters the system with a high velocity. In these cases, however, where the field has opposite directions on the two sides of the aperture, the equipotential distribution is similar rather to the saddle-field type discussed under symmetrical lenses (section 26) than to the simple aperture lens shown in Fig. 37 (*a* and *b*). Attention will be confined here to the two latter arrangements, and we shall assume that electrons originate from the first electrode, which is accordingly the cathode (*C*).

When the first potential gradient is (negatively) greater than the second (Fig. 37 *a*), the equipotentials bulge into the second region and an electron beam is refracted away from the axis; in other words, the lens is divergent when the second derivative of the potential is negative. Conversely (Fig. 37 *b*) the lens is convergent when the second derivative is positive and the equipotentials bulge in the sense opposed to the incident beam. If now the diaphragm is made slightly negative, the positive field will still penetrate to the cathode and extract electrons. In this case the convergent action may be so strong that the beam crosses over the axis within the aperture, forming a 'virtual object' of smaller dimensions than the cathode, and this circumstance is put to use in electron guns (cf. section 47). As will appear later, the divergent system has very limited power; it is important, however, as being the only lens of 'concave' type (or rather, plano-concave) so far constructed.

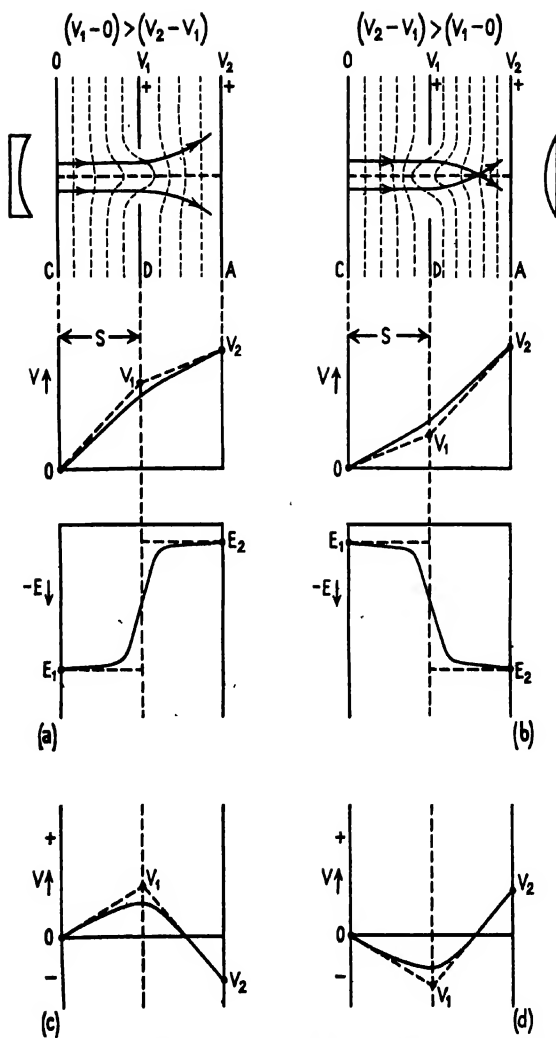


FIG. 37. Field (E) and potential (V) in aperture lenses: (a) negative lens; (b) positive lens; (c) and (d) lens with negative electrode.

Starting from the assumption that the change in field strength takes place entirely within the finite width d of the aperture (Fig. 38), it is possible to derive an expression for the focal length of the lens in terms of the applied potentials. For this purpose it is helpful to transfer attention from the equipotentials to the orthogonal set of lines of force. Assuming the system to be convergent, the field strength will be higher on the side of the diaphragm farther from the cathode; in other words, the lines of force will be closer together there than in the region first traversed by the electron beam. The two field distributions are assumed to be uniform up

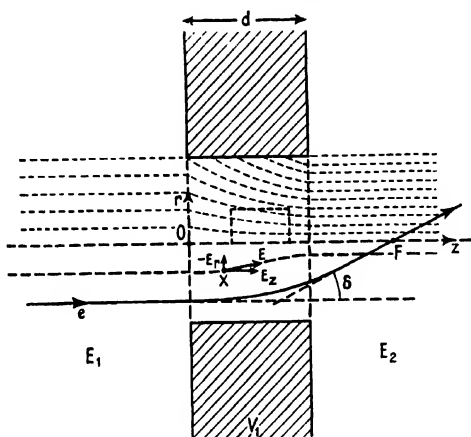


FIG. 38. Aperture lens: transitional field in aperture, and electron trajectory ($E_2 > E_1$).

to the respective surfaces of the diaphragm, the transition taking place within the aperture. Laplace's condition must be obeyed in any small volume element in this region, such as that outlined in the diagram: as many lines of force must leave as enter it. This can only be so if lines pass through the sides of the element, since more enter the far end than leave the near end (with respect to the cathode). The field must, therefore, have both a radial and an axial component within the lens, as indicated in the lower part of the diagram; in any case this follows from the shape of the lines. It is clear also that the radial component, and hence the converging effect on (negative) electrons, must increase with the radius. We have first to find an expression for this radial component and then for the deviation as a function of the radius, in order to evaluate the focal length.

If, at any point X within the aperture, the radial and axial components of the field are E_r and E_z respectively, we may write Laplace's equation in cylindrical coordinates as before:

$$\nabla^2 V = \frac{\partial^2 E_z}{\partial z^2} + \frac{1}{r} \frac{\partial(r E_r)}{\partial r} = 0, \quad (\text{III.10})$$

no term in θ appearing owing to the rotational symmetry of the system.

The total differential of the radial field will be

$$d(rE_r) = \frac{\partial}{\partial z}(rE_r) dz + \frac{\partial}{\partial r}(rE_r) dr. \quad (\text{III.11})$$

We assume, however, that the radial distance of the electron varies only slightly in its passage through the lens and that E_r changes but slowly along its path. This is justifiable if the size of the aperture is small compared with the object-image distance or, in other words, if thin lens and paraxial conditions are assumed. We may then neglect the first term in III.11, and write

$$d(rE_r) = \frac{\partial(rE_r)}{\partial r} dr = -r \frac{\partial E_z}{\partial z} dr,$$

from III.10. Integration gives

$$E_r = -\frac{r}{2} \frac{\partial E_z}{\partial z} = \frac{r}{2} V_z''. \quad (\text{III.12})$$

This expression has already been obtained, with similar approximation, for the two-cylinder lens (equation II.13); a corresponding relation will be obtained later for the short magnetic lens. The sign of E_r is opposite to that of $\partial E_z/\partial z$ since the lines of force draw closer to the axis as z increases.

Suppose now an electron (e) to enter the lens parallel to the axis at a radial distance r_0 . The radial force exerted on it is given by

$$m\ddot{r} = m(dv_r/dt) = -eE_r = \frac{1}{2}er dE_z/dz,$$

or

$$dv_r = (er/2mv_z) dE_z,$$

where $v_z = dz/dt$, the velocity parallel to the axis, and v_r is the radial velocity. Taking the origin at O , in the near face of the aperture, and integrating over its width,

$$\int_{v_{r_0}}^{v_{r_1}} dv_r = (e/2m) \int_{E_{z_0}}^{E_{z_1}} r/v_z dE_z.$$

But, in the thin lens conditions assumed, the value of v_z as well as of r will not change appreciably in the aperture, and we may write

$$\int_{v_{r_0}}^{v_{r_1}} dv_r = (er_0/2mv_z) \int_{E_{z_0}}^{E_{z_1}} dE_z, \quad (\text{III.13})$$

making integration possible. We thus have

$$v_{r_1} - v_{r_0} = (er_0/2mv_z)(E_{z_1} - E_{z_0});$$

but the field is uniform on either side of the aperture, and therefore we may write

$$v_{r_1} - v_{r_0} = (er_0/2mv_z)(E_2 - E_1) \quad (\text{III.14})$$

$$= v_{r_1},$$

since the initial value of the radial velocity was assumed to be zero. Hence the emergent direction of the beam is given by

$$(dr/dz)_1 = v_{r_1}/v_z = (er_0/2mv_z^2)(E_2 - E_1). \quad (\text{III.15})$$

But, if the electrons are emitted from the cathode with zero velocity and accelerated only by the field between it and the diaphragm at potential V_1 , the axial velocity at the aperture is given by

$$mv_s^2 = 2eV_1,$$

and III.13 becomes

$$(dr/dz)_2 = r_0 \left(\frac{E_2 - E_1}{4V_1} \right) = -r_0 \left(\frac{E_1 - E_2}{4V_1} \right) = \tan \delta \approx \delta, \quad (\text{III.16})$$

where δ is the angular deviation suffered by a paraxial beam in the lens. As this deviation proves to be directly proportional to the radial distance (r_0), the primary condition for image formation in a lens system is fulfilled. When the relative value of the fields is such that dr/dz is negative, the emergent beam will be convergent and a real point focus will be formed. The incident beam being parallel to the axis, the position of this point, F , and therefore the focal length f is given by

$$f = r_0/\delta = 4V_1/(E_1 - E_2). \quad (\text{III.17})$$

In the general case, when the object is not at infinity but at a distance p from the lens, equation III.14 becomes, on division by v_s ,

$$\frac{v_{r_2} - v_{r_1}}{v_s} = \frac{er_0}{2mv_s^2}(E_2 - E_1),$$

$$\text{or} \quad \left(\frac{dr}{dz} \right)_2 - \left(\frac{dr}{dz} \right)_1 = \frac{r_0}{4V_1}(E_2 - E_1), \quad (\text{III.18})$$

on eliminating time and writing $V_1 = \frac{1}{2}mv_s^2/e$. As r_0 is assumed constant within the aperture, we may write

$$\left(\frac{dr}{dz} \right)_1 = \frac{r_0}{p}, \quad \left(\frac{dr}{dz} \right)_2 = \frac{r_0}{q},$$

where q is the image distance. Expressions III.17 and III.18 then give at once

$$\frac{r_0}{q} - \frac{r_0}{p} = \frac{r_0}{f}, \quad \text{or} \quad \frac{1}{q} - \frac{1}{p} = \frac{1}{f},$$

the normal optical relation for a thin lens. Hence an isolated aperture has precisely the same image-forming properties for electrons as a material lens has towards light.

In applying the formula for the focal length (III.17) it must be remembered that the field is the negative gradient of the potential, $E = -V'$; so that if the potential rises in the z -direction, the field is negative. Thus, in the arrangement of Fig. 37 *a*, E_1 is numerically greater than E_2 , but both are negative and hence the focal length is negative and the lens divergent. The power of the lens, however, is limited by the circumstance that both f and E_1 depend directly on V_1 , the potential of the diaphragm. Any attempt, therefore, to shorten the focal length by decreasing V_1 will result also in a decrease in E_1 . In fact, the negative power of the lens can

only be increased by increasing V_1 , so that E_1 increases also. Assuming the field to be uniform up to the diaphragm, its value must be

$$E_1 = -(V_1 - 0)/s,$$

where s is the separation of diaphragm and cathode. The focal length will therefore tend to a minimum value of $-4s$ when E_2 is negligible. Hence the power of this 'concave' lens is severely limited.

On the other hand, for the convergent form (Fig. 37 b), the focal length can be decreased without limit by decreasing V_1 , since this will result in a

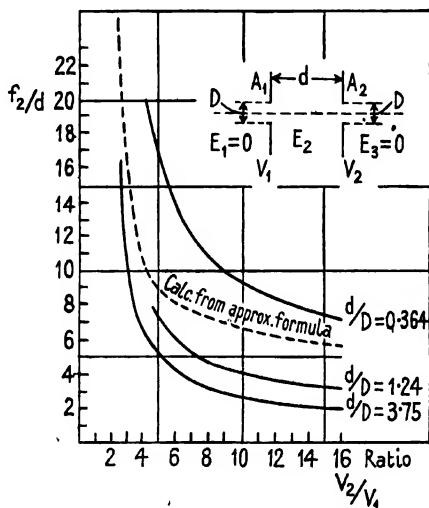


FIG. 39. Focal length of two-aperture lens, for various voltage ratios (V_2/V_1) and separations (d).

corresponding increase in the negative value of E_2 and a decrease in that of E_1 . The power of the lens can thus be very great, and this is the basis of its use in combination with other electrodes in electron guns. It will be appreciated that the above result for the focal length has only been obtained under a number of approximations. However, it proves to be a sufficiently accurate guide for many practical purposes when using aperture lenses; it loses even approximate validity when the width of the aperture is large compared with the separation of the electrodes, or when the field on one side is so great that it penetrates appreciably into the space on the other side. It gives good agreement only for weak fields and long focal lengths.

In the same limiting circumstances it is permissible to use expression III.17 to obtain the resultant power of a combination of two aperture lenses. Suppose two apertures to be set up on the same axis (Fig. 39) at a distance d apart, and potentials V_1 and V_2 to be applied to them. For simplicity we shall assume the field on either side to be zero, $E_1 = E_3 = 0$.

The field between the diaphragms will be given by

$$E_z = \frac{V_1 - V_2}{d},$$

and we shall have for the focal length of each lens separately

$$f_1 = \frac{4V_1}{0 - E_z} = \frac{4V_1 d}{V_2 - V_1}, \quad \text{and} \quad f_2 = \frac{4V_2}{E_z - 0} = \frac{-4V_2 d}{V_2 - V_1}. \quad (\text{III.19})$$

The power of the combination may then be calculated directly from the usual optical expression:

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} + \frac{d}{f_1 f_2}, \quad (\text{III.20})$$

where F is the focal length of the combination.

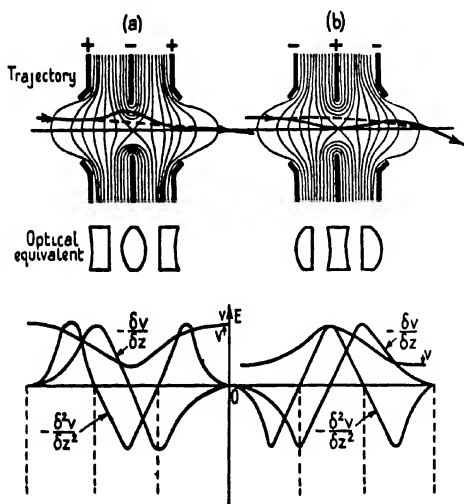


FIG. 40. Symmetrical lens: form of field with central element (a) negative, (b) positive (after Myers).

Measurements of the focal length of such a system have been made by direct experiment by Polotovski, with various potential ratios V_2/V_1 as well as for a number of ratios of aperture separation d to aperture diameter D . The results are shown in Fig. 39, the focal length being given in terms of the separation; the broken line shows the values obtained by calculation from III.19 and III.20. It will be seen that agreement with the simple theory is good for low voltage ratios, and still moderate for ratios as high as 10 or 12 when the separation is not large, that is, in the thin lens conditions for which III.20 is derived.

26. Symmetrical and Univoltage Lenses

The distinguishing characteristic of the symmetrical lens is that the potential is the same on the two sides of the system, or, in other words, the

refractive indices of object and image spaces are equal. A number of possible forms were shown in Fig. 26. Clearly the aperture lens will also become a symmetrical lens when the potentials of the first and last electrodes are made equal; E_1 and E_3 will then be equal but of opposite sign, and the power will not therefore be zero (see III.17). As usually employed, however, all three electrodes have apertures and the conditions are quite different from those in the simple aperture lens. The typical distribution of equipotentials is then as shown in Fig. 40, being entirely symmetrical about the mid-plane of the system. The map is the same whether the central electrode is positive (b) or negative (a) with respect to the outer two, but the variation of potential and of field strength along the axis are reciprocally related in the two cases (lower section of diagram). The equivalent optical systems and typical trajectories through each lens are also indicated. In both cases the total effect of the system is convergent, the positive action of the negative elements prevailing no matter whether the potential difference ($V_1 - V_3$) be positive or negative.

The form of the field in the central aperture is shown on a larger scale in Fig. 41, with the positions of the chief equipotentials marked for the case when this element is negative; when it is positive, the distribution is identical but the value of each line will be found by subtracting the figures given from 100 per cent. The field in the centre of this region can be evaluated from the particular solution of Laplace's equation given earlier (II.11):

$$V_w = \frac{1}{2\pi} \int_0^{2\pi} f(z + ir \sin \alpha) d\alpha,$$

and it has been shown that here the curves are symmetrical hyperbolae. The limiting equipotentials which cross at the centre of the lens make an angle of $54^\circ 44'$ with the axis; in reality, the field in three dimensions corresponds to the rotation of the diagram around the axis, and hence these lines trace out cones with a semi-vertical angle of $54^\circ 44'$. This angle is independent of the shape of the electrodes or their disposition, as has been confirmed by measurement in the electrolytic trough. Away from the centre of the lens, of course, the form of the field varies with the arrangement of the electrodes although the general configuration remains the same. When the central element is at negative potential the field through it resembles a saddle or pass in a mountain range, over which the electron has to travel; when it is positive the field here is more like a 'hog's-back'. On this account, these systems are often referred to as 'saddle-field' lenses. Alternatively, they are known as 'einzeln' (single)

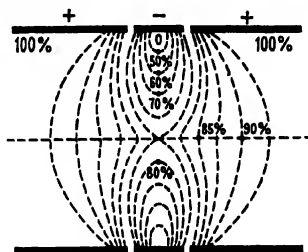


FIG. 41. Symmetrical lens:
detail of the saddle-field.

lenses, since the field in an isolated aperture has this form when the gradient is of opposite sign on the two sides.

We now proceed to obtain a more accurate expression for the focal length than that given by the approximate treatment of the aperture lens: the solution is applicable to all types of electrostatic lens. We proceed from the ray equation (II.16):

$$4Vr'' + 2V'r' + V''r = 0,$$

$$\text{or} \quad -V''r/4 = V'r'/2 + Vr'' = \sqrt{V} \frac{d}{dz} (r'\sqrt{V}), \quad (\text{III.21})$$

where primes indicate derivatives with respect to z , the axial coordinate. Integration gives

$$|r'\sqrt{V}|_a^z = -\frac{1}{4} \int_a^z \frac{rV''}{\sqrt{V}} dz,$$

supposing an electron to enter the lens at $z = a$. If the incident electron be taken as parallel to the axis, in order to find the focus, then

$$(dr/dz)_a = r'_a = 0,$$

and we have

$$r'_z = -\frac{1}{4\sqrt{V}} \int_a^z \frac{rV''}{\sqrt{V}} dz. \quad (\text{III.22})$$

Assuming r to be constant in the lens and integrating, we should at once obtain the approximate expression for the power of an aperture lens (III.16) derived earlier by other reasoning. In the present case, thin lens conditions do not obtain and we cannot assume r to be constant in the lens. It is only possible then to solve equation III.22 by successive approximation. We set

$$r' = r'_0 + r'_1 + r'_2 + r'_3 + \dots = \sum_{n=0}^{n=\infty} r'_n,$$

where

$$r'_0 = r'_a = 0, \quad \text{and} \quad r_0 = r_a.$$

Then

$$\begin{aligned} r'_1 &= -\frac{1}{4\sqrt{V}} \int_a^z \frac{r_0 V''}{\sqrt{V}} dz \\ &= -\frac{r_a}{4\sqrt{V}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} dz, \quad \text{and} \quad r_1 = -\frac{r_a}{4} \int_{-\infty}^z \frac{1}{\sqrt{V}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} dz dz. \\ r'_2 &= -\frac{1}{4\sqrt{V}} \int_a^z \frac{r_1 V''}{\sqrt{V}} dz \\ &= +\frac{r_a}{16\sqrt{V}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} \int_{-\infty}^z \frac{1}{\sqrt{V}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} dz dz dz, \quad \text{and} \quad r_2 = \int_{-\infty}^z r'_2 dz_1, \dots, \text{etc.} \end{aligned}$$

We then have

$$r'_b = r_a \left[\frac{-1}{4\sqrt{V_b}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} dz + \frac{1}{16\sqrt{V_b}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} \int_{-\infty}^z \frac{1}{\sqrt{V_b}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} dz dz dz \dots \right],$$

where r'_b is the total deviation of the beam produced by the lens and V_b is the potential of the image space.

$$\begin{aligned} \therefore 1/f_2 &= -r'_b/r_a \\ &= \frac{1}{4\sqrt{V_b}} \int_{-\infty}^{\infty} \frac{V''}{\sqrt{V}} dz - \frac{1}{16\sqrt{V_b}} \int_{-\infty}^{\infty} \frac{V''}{\sqrt{V}} \int_{-\infty}^z \frac{1}{\sqrt{V_b}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} dz dz dz \\ &\quad + \frac{1}{4^3\sqrt{V_b}} \int_{-\infty}^{\infty} \frac{V''}{\sqrt{V}} \int_{-\infty}^z \frac{1}{\sqrt{V_b}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} \int_{-\infty}^z \frac{1}{\sqrt{V_b}} \int_{-\infty}^z \frac{V''}{\sqrt{V}} dz dz dz dz dz \dots \end{aligned} \quad (\text{III.23})$$

A numerical evaluation of this expression may be made once the axial distribution of potential and its derivatives have been found.†

As the electron has been considered to be travelling in the z -direction, this result gives the second focal length of the lens. The first focal length may then be calculated if the initial and final voltages are known. From III.21 we have simultaneous differential equations for two particular ray-paths through the system, designated r_α and r_β :

$$\begin{aligned} \sqrt{V_z} d/dz(\sqrt{V_z} r'_\alpha) &= -V''_z r_\alpha/4, \\ \sqrt{V_z} d/dz(\sqrt{V_z} r'_\beta) &= -V''_z r_\beta/4. \end{aligned}$$

Multiplying by r_β and r_α respectively, and subtracting, we have

$$d/dz(r_\beta \sqrt{V_z} r'_\alpha - r_\alpha \sqrt{V_z} r'_\beta) = 0,$$

whence

$$r_\beta r'_\alpha - r_\alpha r'_\beta = C_{\alpha\beta}/\sqrt{V_z}, \quad (\text{III.24})$$

where $C_{\alpha\beta}$ is a constant of integration depending on the two rays chosen. They are thus interrelated, as regards distance from the axis, and this relation holds for every value of z ; it is equivalent to the well-known rule of optics that two rays suffice to describe completely the focusing properties of a system. Any other ray path may be expressed in terms of two such special solutions (see p. 109); here we employ them to obtain the relation between the two focal lengths. It is especially to be noted that r_α and r_β here stand for the equations of two separate rays, and not for two values of the radial ordinate on a given ray, as did r_a and r_b above.

In the object space, where $V_s = V_a$, it is convenient to choose the path of r_α as passing through the first focal point, at z_f , with inclination α to the axis; thus

$$r'_{\alpha(o)} = \tan \theta_1 = \alpha,$$

† After setting $R = rV^{\frac{1}{2}}$ in the ray equation (III.21), Scherzer obtains a similar expression for f , but containing the first instead of the second axial derivative of the potential; this is advantageous in that V' can be experimentally determined with much greater accuracy than can V'' (cf. also p. 111).

and the equation of the path is

$$r_{\alpha(o)} = a(z - z_{f_1}).$$

In order to locate the second focal point we must now allow r_β to be parallel to the axis, at a distance b from it, in the object space, so that

$$r_{\beta(o)} = b, \quad r'_{\beta(o)} = 0.$$

Substituting in III.24 we then have

$$C_{\alpha\beta} = ab\sqrt{V_a}, \quad (\text{III.25})$$

and therefore

$$r_\beta r'_\alpha - r_\alpha r'_\beta = ab\sqrt{(V_a/V_s)}. \quad (\text{III.26})$$

Now in the image space it follows that r_α will be parallel to the axis (at some distance c from it), and r_β will pass through the second focal point, at z_{f_2} , with inclination $-d$:

$$r_{\alpha(i)} = c; \quad r'_{\alpha(i)} = 0; \quad r_{\beta(i)} = d(z_{f_2} - z); \quad r'_{\beta(i)} = \tan \theta_2 = -d.$$

On substitution in III.26 we get

$$-cd = ab\sqrt{(V_a/V_s)}. \quad (\text{III.27})$$

But the focal lengths are given by

$$\begin{aligned} c/f_1 &= \tan \theta_1 = a, \\ b/f_2 &= \tan \theta_2 = -d, \end{aligned}$$

whence

$$-cd/f_1 = ab/f_2.$$

Substituting in III.27: $f_1 = -f_2\sqrt{(V_a/V_s)},$ (III.28)

since in the image space the potential $V_s = V_b$. If one focal length is known, therefore, the other may be calculated from a knowledge of the over-all potential ratio applied to the lens. (Cf. III.6.)

Sometimes it is possible to fit the axial distribution of potential in a lens to some simple function, with sufficient accuracy, and then to carry out directly the integration of the expression for the focal length (III.23). For instance, it was indicated by Scherzer that in the symmetrical lens the axial field could be represented by an expression of the form

$$V_s = V_0 + Ae^{-be^2}. \quad (\text{III.29})$$

Pläss has taken the simple distribution

$$V_s = V_0(1 - \frac{1}{2}e^{-is^2}),$$

and evaluated the focal length of the lens from III.23. He found that the series converged rapidly, his results being:

$$\begin{aligned} &\text{from the first term only, } f = 21.32 \text{ cm.} \\ &\text{,, ,, two terms, } f = 12.97 \text{ ,,} \\ &\text{,, ,, three terms, } f = 12.42 \text{ ,,} \\ &\text{,, ,, four terms, } f = 12.36 \text{ ,,} \end{aligned}$$

By plotting the trajectory a value of 13.35 cm. was obtained. For most practical purposes, therefore, it is sufficient to take the first two terms.

Earlier workers had relied on the use of the first term of III.23 only, obtaining by partial integration of the corresponding expression for the first focal length

$$\begin{aligned} 1/f_1 &= \frac{-1}{4\sqrt{V_a}} \left[\left| \frac{V'}{\sqrt{V}} \right|_{-\infty}^{\infty} + \frac{1}{2} \int_{-\infty}^{\infty} \frac{V'^2}{V^{\frac{3}{2}}} dz \right] \\ &= \frac{-1}{8\sqrt{V_a}} \int_{-\infty}^{\infty} \frac{V'^2}{V^{\frac{3}{2}}} dz, \end{aligned} \quad (\text{III.30})$$

since V' is zero at the limits of integration on either side of the lens in the case of two-cylinder and symmetrical systems; V_a is the potential of the object space.

The distribution of potential along the axis of the two-cylinder lens is closely reproduced by the function

$$V_{0z} = V_0 + A \tanh z.$$

Plass calculated the focal length from III.23 using the distribution $V_{0z} = V_0(1 - \frac{1}{2} \tanh z)$, and a total (decelerating) voltage ratio of 3:1. He obtained:

$$\begin{array}{lll} \text{from the first term only,} & f = 15.85 \text{ cm.} \\ \text{,, ,, two terms,} & f = 10.75 \text{ ,,} \\ \text{,, ,, three terms,} & f = 10.85 \text{ ,,} \end{array}$$

Ray-tracing gave a value of 10.81 cm.; clearly here also it is not accurate enough to take the first approximation, but two terms will suffice for most purposes.

Direct experiments on the form of symmetrical lens shown in Fig. 40 had previously been carried out by Johannson and Scherzer, the potential V_1 of the outer electrodes being maintained constant at 750 volts whilst V_2 was varied. The observed focal length (f) is plotted against the ratio $(V_2 - V_1)/V_1$ in Fig. 42. In interpreting the graph it must be remembered that when $V_2 - V_1 = 0$, all three electrodes are at 750 V. with respect to the cathode. It will be seen that the focusing action is strongest when the middle electrode is negative with respect to the other two: that is, when there is a potential trough at the top of the gradient up which the electrons mount to the centre of the lens. Owing to the retarding field, the electrons have low velocity in this region and the effect on them of the converging field is correspondingly large.

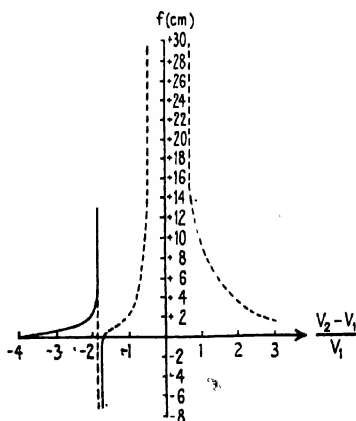


FIG. 42. Symmetrical lens: variation of focal length with voltage ratio and conversion to mirror (full line).

When the potential V_2 is made equal to that of the cathode from which the electrons originate, only one voltage need be applied to the system, and $V_2 = 0$. Since the velocity of emission of electrons from the cathode is negligible, the ratio V_1/V_2 is now infinite for all values of V_1 , and the focal length of the lens is constant and very short. Its power, therefore, is independent of fluctuations in the voltage supply, and 'chromatic aberration', which otherwise arises from irregular changes in the velocity of the electrons accelerated from the cathode, is automatically eliminated. This arrangement has great practical advantages and is much used; the lens is then styled a 'univoltage lens'. It is found, however, that its focal length depends very critically on the length of the central electrode, or tube, as it frequently may be. As the length increases, for a given diameter, the focal length diminishes rapidly towards a limiting value for lengths greater than twice the radius. This change is associated with a change in the distribution of the field near the saddle-point of the lens: the value, as a percentage of V_1 , of the equipotential which crosses the axis at this point falls sharply as the length increases. Hence the potential at the saddle-point approaches closer to zero, and its ratio to V_1 accordingly increases; the focal length is thus shortened, since it is the ratio of the potentials on the axis that determines the focusing power for a paraxial beam. For the same reason, the voltage of the middle electrode can be decreased appreciably below that of the original cathode (-750 V. in terms of Fig. 42) before electrons are entirely prevented from passing through the lens, and this lag is greater the shorter the electrode, since the potential trough at its centre is then deeper relative to the potential at its circumference.

The symmetrical lens is usually used in this univoltage form and finds application in cathode-ray tubes, the electrostatic electron microscope, and many other devices. It is possible, of course, to destroy the symmetry by applying different potentials to the two outer electrodes and still have a practicable lens. The saddle-point then moves away from the mid-plane in the direction of the greater potential gradient. The focusing action of the system remains essentially the same as in the symmetrical case, but it is not in general use owing to the obvious advantages of the latter.

27. *Immersion Lenses*

In the optical sense the two-cylinder lens is properly of the immersion type, since object and image space differ in potential and hence in refractive index. It is convenient in electron optics, however, to reserve the title to those lenses in which the object (usually the cathode itself) is deeply immersed in the field, so that the refractive index varies rapidly in its neighbourhood. It follows that electrons enter the lens with almost zero velocity; they are in fact extracted from the cathode under the influence of the applied potentials, and have initially only the velocity of thermal emission. These conditions do not obtain in the normal two-cylinder lens, but are naturally fulfilled in almost all electron guns. It should be noted, however, that some authors include both types under

the heading of immersion lenses, and apply the name of 'immersion objective' to distinguish the type discussed in this section.

A number of forms of the immersion lens were shown in Fig. 27. Usually it consists of two apertures, with or without attached tubes, arranged close to the cathode (Fig. 43). The first is often, but not necessarily, at negative potential relative to the cathode (C) and is known as the grid (G); the second is at positive potential, in order to accelerate the electrons, and is the anode (A). The form of the equipotentials and typical trajectories are as shown: the lens is convexo-concave and usually of such high power that the beam crosses over the axis within the lens itself, and therefore leaves it divergent.

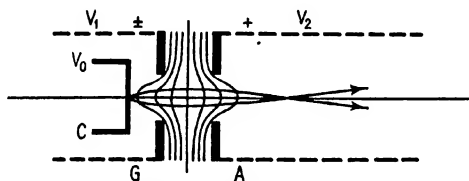


FIG. 43. Three-element immersion lens.

The simplest immersion lens would consist of a single aperture at positive potential, to accelerate the electrons. But it follows from the treatment of aperture lenses (Fig. 37 *a*) that this arrangement could not have a converging action in the presence of a plane cathode. The necessary convex curvature can be given to the equipotentials either by using a concave surface for the cathode, or by introducing a second diaphragm at higher potential. The first diaphragm may remain positive, may be at cathode potential (Wehnelt cylinder), or may be made negative; the essential requirement is that the potential of the second (= anode) be high enough to create a greater field between it and the first than that between the latter (= grid) and the cathode (cf. section 25, on aperture lenses). It is found in practice that the power of the lens is increased if the grid is made slightly negative, providing that the positive field of the anode can still penetrate beyond it to the cathode. This condition is achieved by having a slightly larger hole in the grid than in the anode and by arranging the two electrodes very close together. (For the form of the field between cathode and anode in the presence of a cylindrical grid, see reference to Grösser on p. 76.)

The cathode being fixed at zero potential, the focal length of the immersion lens depends on the ratio of the grid and anode potentials, V_1/V_2 . The nature of this dependence has been investigated by Johansson and others, and has the form indicated in Fig. 44. The focal length decreases, and the magnification increases, as the grid voltage is reduced and finally made negative. The minimum focal length is about 1.5 times the diameter of the grid aperture. By reducing the size of this aperture, magnifications greater than 200 may be obtained. The system is therefore useful as a low-power

electron microscope for examining the surface of the emitting cathode. In the cathode-ray tube, on the other hand, the magnification must be kept low in order to give a small spot on the fluorescent screen. This may be effected by reducing the voltage ratio, and by increasing the diameter of the grid aperture and the separation of the grid and cathode. Advantage can also be taken of the cross-over of the beam inside the lens (Figs. 43 and 45) as this has a smaller cross-sectional area than the cathode itself. Auxiliary anodes are then arranged to act as a second lens which projects an image of this cross-over disk on the viewing screen, of smaller size than could be obtained from the cathode itself.

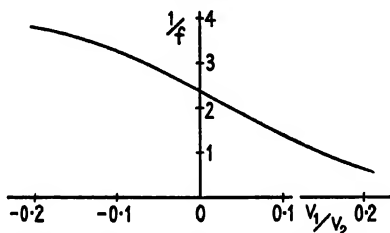


FIG. 44. Three-element immersion lens: variation in focal length with voltage ratio (V_1/V_2).

The arrangements described so far form the basis of practical electron guns in circumstances where an intense beam is required to be focused into a spot; further details will be given in the discussion of electron optical applications. It sometimes happens, however, as in television transmitting cameras, that a lens is needed to reproduce a true image of an extended cathode with little or no magnification. This demand was met by Zworykin by adding a cylindrical tube to the cathode, and thus creating in effect a two-cylinder lens between it and the anode, with the second tube usually the longer (Fig. 45). The cathode, which is now a photo-electric surface illuminated on the side away from the anode, must lie just within the field of the anode in order that adequate emission shall occur, that is, the object is within the focusing field. Ideally, V_0 should be zero, in which case the refractive index of the object space is zero and $f_1 = 0$, the first principal plane being at the object. The system thus differs from the simple two-cylinder lens in that the initial velocity of the incident electrons is effectively zero. It is clear, nevertheless, that the focusing action is essentially the same as in the parent lens, the distribution of the equipotentials and the axial field being only slightly disturbed by the presence of the cathode. The analytical treatment, due to Morton and Ramberg, follows the lines indicated for the two-cylinder lens, and yields a solution in similar form to II.5. In cylindrical coordinates the axial potential distribution is given by

$$V_{(z)} = \frac{2}{\pi} \int_0^{\infty} V_1 \frac{\cos ku \sin kz}{k J_0(ik)} dk,$$

from which V'_{0a} and V''_{0a} are immediately obtained; as before, these expressions have to be evaluated by quadrature. The form of V_{0a} and V'_{0a} is shown in Fig. 45, together with four typical trajectories. The trajectories depend only on u , the distance of the cathode from the gap between the cylinders (= object distance), and are independent of the anode potential V_1 . Consequently it is impossible to vary the focal length in a system of given dimensions. To overcome this difficulty a grid cylinder may be introduced, or alternatively the cathode cylinder is split up into several

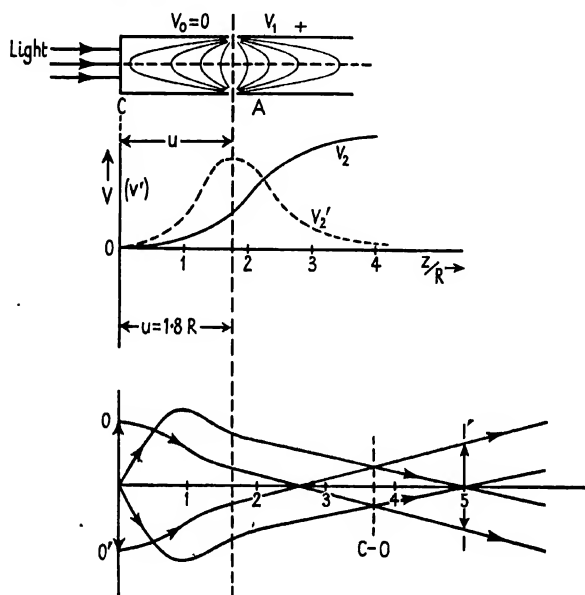


FIG. 45. Image-tube immersion lens: potential distribution and trajectories (Zworykin).

separate ring elements at successively higher potential as described in Chapter VI (section 51). It is usual also to curve the cathode surface in order to minimize aberrations of the image, which, as indicated by the trajectories, is produced immediately beyond the end of the anode cylinder. The image falls on a secondary emitting surface, which yields several electrons for each that is incident. The system thus constitutes an amplifier, or multiplier-tube, and is referred to as an 'image-converter', since the optical image first formed on the cathode surface is faithfully reproduced as an electron image.

28. Electron Mirrors

Any of the lens systems discussed above can be converted into an electron mirror by setting the negative potential of one of the electrodes at a lower level than that of the emitting cathode. In electron optics, therefore, in contradistinction to light optics, it is convenient to consider

mirrors as a special case of lens arrangements. They have not so far attracted as much attention as electron lenses, on account of the inherent difficulties of practical application; it is less easy to extract an image from the system than in ordinary optics. The properties of the symmetrical and the two-cylinder types have been investigated, however, and it has been shown that both convex and concave forms can be produced. The aperture and immersion lenses lend themselves less readily to adaptation as mirrors.

The Two-cylinder Mirror. The behaviour of the two-cylinder system as a mirror has been investigated by Nicoll. He used a first lens to project a primary image into the mirror, through a hole in the screen on which the final image was viewed. The arrangement of the apparatus is shown in

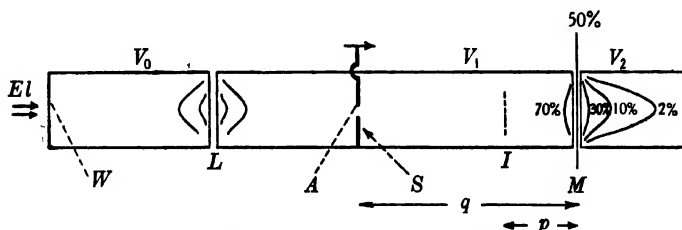


FIG. 46. Two-cylinder electron mirror: Nicoll's experimental arrangement.

Fig. 46, where L is the first lens and M the mirror formed between tubes of equal radius R . Electrons from an electron gun enter the system through a wire gauze W , which acts as primary object and facilitates the measurement of magnification. The lens L forms an image, through the aperture A , in some plane I , the position of which may be calculated from a knowledge of the cardinal points of L for different applied voltage ratios $V_1:V_0$. The final electrode may then be reduced to such a negative potential $-V_2$ that the electron beam is reflected back on to the screen; the first image I acts as object for the mirror and produces a final sharp image on S when the distances and voltage ratios are properly adjusted. Reflection will take place at that equipotential surface which is at zero potential with respect to the incident electrons; their forward velocity is reduced to zero by the time they reach this position, and they are then accelerated in the opposite direction by the positive field of V_1 . If the original accelerating potential applied to the beam is also V_1 , then this 'reflection' will occur at a surface which is at a potential given by

$$V_r = \frac{(V_1 + V_2) - V_1}{V_1 + V_2} = \frac{V_2}{V_1 + V_2},$$

as a fraction of the total potential difference between the two electrodes forming the mirror. When the voltage ratio $V_1:V_2$ is unity, this reflecting equipotential will be that at one-half (or 50 per cent.) of the total difference. As the ratio is increased, reflection will occur at surfaces of lesser denomination; for a ratio of 4:1 it will take place at the 20 per cent. equi-

potential. As in all two-cylinder lenses, the distribution of equipotentials is symmetrical about the mid-plane M , as is again depicted here.

The procedure followed by Nicoll was to fix the voltage ratio applied to the mirror (V_1/V_2) at a number of successive values and to find for each the value of the ratio across the lens (V_1/V_0) needed to produce a clear image of the gauze W on the screen S , which was kept at a fixed distance ($q = 3R$) from the mirror. The position of I was then calculated in each case. In other words, the first object W and the final position of the image

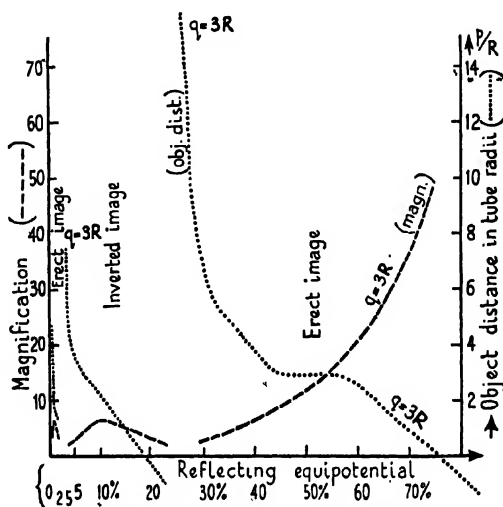


FIG. 47. Magnification in two-cylinder electron mirror (with fixed image distance). (Nicoll.)

(at S) being fixed, the position of the first image I was varied with respect to the fixed mirror M , it being more convenient to vary the power of the lens L than to move the viewing screen. The latter was observed through the end of the apparatus beyond M .

In this way the object: image relationship was investigated as a function of the voltage ratio across the mirror, and the results are given in Fig. 47. It is helpful to plot the object-mirror distance ($IM = p$), as well as the magnification, against the voltage ratio; the value of the reflecting equipotential V_r is also given. When the ratio is 1:3, giving a strong reflecting field, the mirror acts as a convex surface, reducing the convergence of the incident beam and forming an erect image. Reflection takes place at $V_r = 75$ per cent. above the potential of the final electrode, of a virtual object located almost at the mirror; the magnification is correspondingly large. As the ratio is increased the reflecting equipotential decreases in relative value and is located farther to the right until at 1:1 it is in the mid-plane; the object distance increases and the magnification decreases. Beyond a ratio of 2:1 the object distance moves away to infinity, and then

assumes negative values. The image now proves to be inverted and the magnification small; the system is acting as a concave mirror for a 'real' position of I . The object distance, however, rapidly increases again with increasing ratio until the positive field between M and S becomes so powerful that it converges the beam again after a first cross-over, producing an intermediate image after I ; a second image is then focused on the screen. This takes place at a voltage ratio of about 20:1, when reflection occurs at the 5 per cent. equipotential, as is indicated to the left of Fig. 47. Higher values of $V_1:V_2$ will result in the formation of further intermediate images, the beam crossing the axis several times. In these circumstances the beam penetrates deep into the final electrode before suffering reflection, and the aberration effects are increased. There is a tendency also for space charge to accumulate in the negative electrode, and finally no clear image can be formed at all.

For lower values of the voltage ratio, however, the electron mirror is found to give images of moderately good definition. As will be shown later (cf. p. 125), its spherical aberration is of negative sign, whereas all lenses have positive values. The possibility exists, therefore, of using it to correct lens defects; but owing to the difficulty of projecting the final image very little progress has been made in applying it to practical devices.

The Symmetrical Mirror. Similar reflecting conditions can be produced in all the other electrode arrangements which normally act as lenses by lowering the potential of one of the later tubes or apertures below that of the original cathode. It is also possible to fashion the negative element as a curved surface so that it not only produces a correct curvature of the equipotentials but bears a physical resemblance to an ordinary optical mirror. The system which has received most attention, however, is that of the symmetrical lens with the central element negative (Fig. 40 a).

The mirror action of the symmetrical field may be best understood, in detail, by considering first the electron path through it when the potential of the central element is made equal to that of the origin of the incident beam. The potential on the axis at the saddle-point will then be slightly higher than that of the surrounding electrode, so that electrons are still able to pass through the system, which will still act as a positive lens. When passing through the diverging fields of the two positive apertures the electron will have high velocity and will suffer only small deviation from the axis; although the radial acceleration may be large, the extent of the deviation is determined by the time of passage in the field. In mounting the potential hill formed by the central electrode, however, the electron is decelerated and finally passes across the saddle with very small velocity. Under these circumstances the converging action of the (negative) potential trough in the centre of the lens is very pronounced, and the electron suffers strong deviation towards the axis. This positive action always outweighs the negative action of the outer electrodes, and is most marked near the critical negative potential, as is evident from the short focal length in the top left quadrant of Fig. 42.

If the saddle potential is now made slightly more negative, electrons will fail altogether to pass across it. They will enter the converging field just below the mid-point and will be directed towards the axis. In the critical condition they will oscillate to and fro across the axis in the potential trough; those with velocity above the average may eventually cross to the other side (path 1, Fig. 48), and those with smaller velocities will be reflected as from a concave mirror (paths 2 and 3). Both the accumulation of space charge at the centre, and the negative action of the

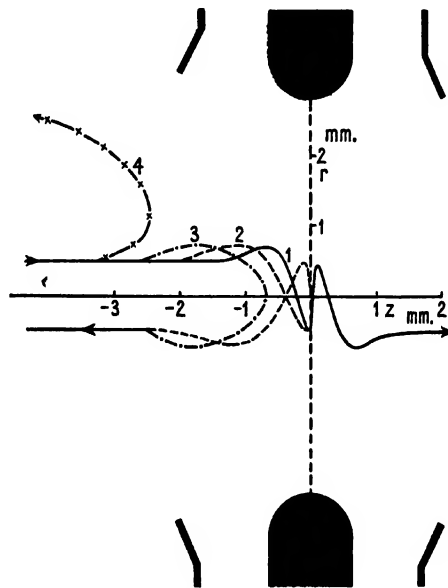


FIG. 48. Trajectories in symmetrical lens—mirror (after Myers).

positive electrode through which these electrons must again pass, will assist this mirror effect. The focal length of the system is accordingly negative, as shown to the left of Fig. 42, and increases rapidly to infinity as the negative potential is further increased. For greater negative ratios it assumes a positive sign which falls to very small values beyond a ratio of about $-2:1$. Under the influence of such a large decelerating potential electrons now do not succeed in rising high enough on the potential hill to reach the trough near the centre; they are turned back when still in a region where the radial acceleration is directed away from the axis. Hence they do not cross the axis at all and are reflected from the system as a strongly divergent beam; the action is now that of a convex mirror, and a typical path is shown (4) in Fig. 48.

The symmetrical system can therefore be used as a positive or negative mirror, by variation of the potential of the central electrode. It would also be practicable to use it as a velocity filter when an incident beam of

electrons was not homogeneous, but contained a straggle of velocities below the maximum value. On the other hand, when the beam is already monochromatic, it could be used to modulate its intensity by causing the potential of the negative element to fluctuate about its critical value: as the negative potential increased, the marginal electrons would be reflected first, then more and more of the beam, until finally the paraxial rays were also turned back. It must be said, however, that the practical applications of such arrangements have so far been very few in comparison with the many uses of lenses.

FURTHER READING

1. Location of Cardinal Points:

Epstein, *P.I.R.E.* **24**, 1095, 1936.

Klemperer and Wright, *Proc. Phys. Soc.* **51**, 296, 1939.

2. Two-cylinder Lenses:

Maloff and Epstein, *Electron Optics in Television*, McGraw-Hill, 1939.

Klemperer and Wright, loc. cit.

Spangenberg and Field, *Elect. Commun.* **21**, 194, 1943.

3. Aperture Lenses:

Davission and Calbick, *Phys. Rev.* **38**, 585, 1931, and **42**, 580, 1932.

Bedford, *Proc. Phys. Soc.* **46**, 882, 1934.

4. Symmetrical Lenses:

Johannson and Scherzer, *Z. Phys.* **80**, 183, 1933.

Plass, *J. Appl. Phys.* **13**, 49, 1942.

5. Immersion Lenses:

Johannson, *Ann. Phys.* **18**, 385, 1933, and **21**, 274, 1935.

Morton and Ramberg, *Physics*, **7**, 451, 1936.

Grösser, *Beiträge zur Elektronenoptik*, p. 82 (1937).

6. Electron Mirrors:

Hottenroth, *Z. Phys.* **103**, 460, 1936; *Ann. Phys.* **30**, 689, 1937.

Recknagel, *Z. Phys.* **104**, 381, 1937.

Nicoll, *Proc. Phys. Soc.* **50**, 888, 1938.

CHAPTER IV

MAGNETIC FOCUSING

29. *Introductory*

THE deviation of electrons in a transverse magnetic field has been discussed in Chapter I; in a uniform field a homogeneous beam is bent into a circular arc or, in the limit, into a complete circle, but there is no true focusing effect (see, however, section 64). However, it was early discovered by Wiechert (1899) that a short current-carrying coil had a concentrating action on electrons passing axially through it. Although the reason for this behaviour in a longitudinal magnetic field was not understood, such concentrating coils were in continuous use in the succeeding years; their action was later strengthened by the addition of a soft-iron housing around the windings. It was not until 1926 that Busch showed analytically that such a system would have the properties of a lens, and derived an expression for the focal length in terms of the field strength and the electron velocity. Since then magnetic lenses have been thoroughly investigated both theoretically and practically. The short lens (Fig. 50) in which image and object are well outside the field has been put to many uses; in certain respects it is analogous to the symmetrical electrostatic lens, and always has a positive power. A special form of it has been developed as an immersion objective (Fig. 56), of focal length comparable with the best light optical objectives; with its aid electron microscopes have been built to give a resolving power two orders better than that achievable with light. On the other hand, solenoidal or long fields (Fig. 49) are also used for some purposes where focusing without magnification is required; as object and image are then in regions of the same refractive index, this system bears some resemblance to the electrostatic symmetrical lens, but is more akin to the image converter immersion lens described in section 27. Magnetic electron mirrors are not known. In respect of theoretical discussion less attention has been given to magnetic lenses than to the electrostatic types, especially in the development of ray-tracing methods and the treatment of aberrations. In one special case, however, it has been possible to obtain an exact expression for the focal length of a magnetic lens (cf. p. 91).

30. *Solenoidal Lenses*

We commence discussion of the focusing action of magnetic fields with the consideration of solenoidal systems (Fig. 49), in which the path of an electron lies entirely within a uniform longitudinal field. It is questionable whether the name of lens is properly applicable to such an arrangement, since it is in theory of infinite length and hence constitutes a different problem from that of thick lenses in light optics. It is convenient, however,

to include it among other electron lenses, especially as the action of a 'thin' magnetic lens can be understood more easily in the light of a discussion of the long field.

Suppose a point source of electrons O to be situated in a uniform longitudinal magnetic field of strength H generated by a solenoidal winding (shown by hatched lines), and electrons to be emitted with velocity v in a direction making an angle θ with that of the field. We may discuss the resulting motion in terms of the two components of velocity, $v \sin \theta (= v_y)$ and $v \cos \theta (= v_z)$, perpendicular to and along the field respectively. v_z will be uninfluenced by the field and the electron will thus have constant velocity in the z -direction, so long as no electrostatic field is applied. In its motion at right angles to the field, however, it will be

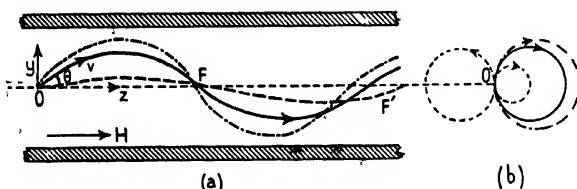


Fig. 49. Trajectories in a solenoidal magnetic field: (a) in a rotating meridional plane; (b) in projection on a transverse plane.

under the influence of a magnetic force (eHv_y) normal to the plane of reference; in the conditions of Fig. 49 this will act out from the diagram. On being deflected in this direction the electron will still be moving normal to the lines of force and will continue to experience a magnetic force which is at all points normal to its direction of motion. As already indicated in section 3, the resulting path, in the case when $v_z = 0$ (or $\theta = 90^\circ$) will be a circle of radius r given by

$$r = mv_y / He \quad (\text{IV.1})$$

in the plane through O normal to the field. In the present case, $0 < \theta < 90^\circ$, the path will be traced out on the surface of a cylinder of the same radius, the electron recrossing the z -axis at an interval determined by the value of v_z ; in other words, it will be a helix with the z -axis tangential to each loop. The projection of the path on to the $x-y$ plane (Fig. 49 b) will be a circle of radius given by IV.1, and in a plane that rotates about the z -axis with one-half the angular velocity of the electron the path will be as shown in Fig. 49 a (cf. also Fig. 4). Such a motion of the electron occurs in all magnetic lenses, resulting in a rotation of the image with respect to the object, and it is convenient therefore to use this method of depicting the path in two planes, one of which rotates with the electron. It needs to be remarked that, for an electron projected with a velocity component in the (positive) y -direction, this rotation is always clockwise when viewed along the direction of the field, as follows from the Left-hand Rule; but it will give a right- or left-handed circular path, according to the direction of the field, when viewed from the point of

origin of the electron. Consequently, if electrons are projected from F in the direction of those arriving from O (one turn of the spiral will have been completed by this time, electrons will arrive from the *negative y*-direction, and Fig. 49 *a* will have become inverted), then they will travel back to O with an opposite rotation, as indicated by the dotted circle to the left of the projection on the x - y plane. The trajectories of electrons in magnetic fields are, therefore, not reversible; the optical principle of reversibility of path is not valid in such lenses.

The time (t) required for the electron to make one complete revolution and return to the x -axis is at once given by

$$t = 2\pi r/v_y = 2\pi m/He, \quad (\text{IV.2})$$

and proves to be independent of both the initial velocity of projection and the angle θ . Electrons of whatever velocity will all have the same period of revolution in a constant field, and it is on this circumstance that the operation of the cyclotron and similar devices depends. On the other hand, the distance z_f travelled longitudinally before an electron again crosses the axis is

$$z_f = v_z t = u \cos \theta (2\pi m/He), \quad (\text{IV.3})$$

and hence will only be the same for electrons projected with identical velocities at the same angle to the field. When θ is small, however, and $\cos \theta$ can be taken as unity for all electrons, then the axial length of the spiral path will be effectively the same for each, and we can speak of a focus at F , the point where they again meet the axis, as shown by the full-line paths of differing radii in Fig. 49. A paraxial and homogeneous beam starting from O will meet again in an approximate point focus at F , where OF has the value z_f given by IV.3; there will, in fact, be a series of focal points equally spaced in this way, at each successive intersection of the helix with the axis. An extended object will give an image of equal size, the magnification being unity; the image will be erect, since a full 360° rotation is experienced. However, an electron beam entering initially parallel to the axis will obviously be under no magnetic influence and will not be focused at all. Strictly speaking, therefore, the focal length of the solenoidal field is infinite, and the ordinary optical relations cannot be applied. The fundamental reason for the special properties of such 'lenses' lies in the fact, which follows from the expressions given above, that the refractive index of the field for the beam is not a function of position only but depends also on the instantaneous direction of the beam. The problem is essentially the same as that of the passage of a beam of light through an anisotropic crystal, and Glaser has treated it from this point of view.

When electrons are emitted from O with differing velocities, it follows from IV.3 that there will be a different point of intersection with the axis for each velocity of emission; the higher the velocity the greater the pitch of the helical path traversed, and the greater the object-image distance.

It thus becomes possible to separate the initial beam into its velocity components by placing limiting diaphragms, or the aperture of a recording device, at selected positions along the axis. This method of 'focal isolation' has been used for β -ray spectrometry by a number of workers using a radioactive source at O and a particle-counter or Wilson chamber at F . It was first suggested by Kapitza, and applied by Tricker; lately it has been discussed and improved by Witcher. The accuracy of resolution is increased by using a line source and a line slit at F , since the paths travelled by electrons of unwanted velocities will at this point have suffered rotations of more or less than the full 360° for which the slit is adjusted.

A second application of solenoidal focusing, or the 'long' magnetic lens as it is frequently called, has been in the construction of image-converters in television, where electrons released photo-electrically have to be brought to a true image on an anode. Zworykin's electrostatic image-tube was described in section 27; Farnsworth has made use of a solenoidal magnetic field for the same purpose. The point source O of Fig. 49 is now replaced by a photo-sensitive cathode and the collecting anode is a plate fixed at F . In order to get a clear image it is necessary to accelerate the electrons to a high velocity by maintaining a potential difference of 700 volts between cathode and anode; to ensure that the electrostatic field is strictly uniform along the tube, the voltage is applied at either end of a high resistance formed by a thin metallic film sputtered on to the glass walls between anode and cathode. The Farnsworth image-transformer, therefore, is a combined magnetic and electrostatic focusing system. It is discussed more fully in Chapter VI (Fig. 97).

31. *The Short Magnetic Lens*

A short magnetic field differs fundamentally in its focusing action from the solenoidal field. Not only are both object and image generally outside the field, but the deviation of the electrons is now mainly the work of the radial component of the field and not of its longitudinal component. In some respects the short magnetic lens is similar to the electrostatic symmetrical lens (cf. Figs. 41 and 55), but the path of the electron in it is complicated by the rotation that occurs about the axis. In order to make this clear, in now deriving an expression for the focal length of the lens, we shall make use of a simplified procedure due to Bouwers. The rigorous treatment is given in section 36, for combined electric and magnetic fields.

If a short homogeneous magnetic field could be produced so that it had abrupt limits, it would have a similar focusing action to the solenoid; the deviation suffered by electrons would depend on their initial direction as well as on their velocity ($=$ wave-length), and no unique focal point would be obtained. In practice the field must have a more or less gradual beginning and end, and it is upon the radial component thus introduced that the focusing effect mainly depends. As shown in Fig. 50, the field due to a short coil will be constant over only a very small region at its

centre, on either side of which the lines of force diverge rapidly from the axis. In such a field the path of the electron will be as represented by the full lines, where (b) is the projection of its motion on a plane at right angles to the axis zz' , and (a) shows its motion in a meridional plane that rotates with the electron about the axis at an angular velocity that is not in general constant throughout the lens. If the electron is projected from a point O on the axis, at an angle θ to it, the path will be straight (in the absence of stray fields) until it enters the magnetic field. The field falls off rapidly along the axis on either side of the equatorial plane PP' and is not necessarily symmetrical about it. Especially when the generating coil is shrouded with soft iron (heavy lines), the effective magnetic field is concentrated in a very small region, and the radial component H_r is

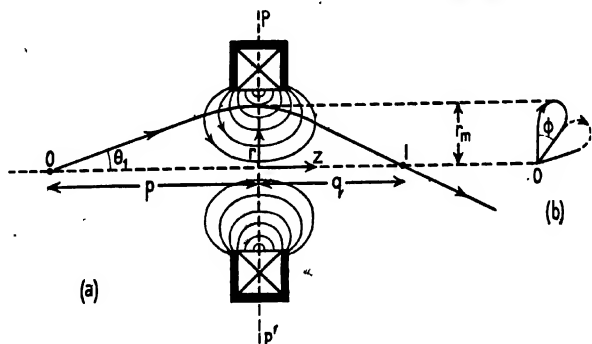


FIG. 50. Trajectory in short magnetic lens: (a) in a rotating meridional plane; (b) in projection on a transverse plane.

correspondingly great. Confining attention to paraxial rays, the value of the radial component of velocity $v_r (= v \sin \theta)$ will be small when the electron enters the magnetic field, and will decrease during its passage through the lens, on account of the acceleration of the electron towards the axis. It follows that this slight radial motion of the electron at right angles to the longitudinal component of the field (H_z) only gives rise to a small force upon the electron; in the first-order theory of the short lens this interaction, which is the determining feature in the solenoidal case, is left out of account.

The focusing action of the lens may be considered to be due to the influence of the radial component of the field on the longitudinal motion of the electron. In the first half of the lens the radial component is assumed to be directed towards the axis, and by the simple rules of electromagnetism, the electron will be deflected upwards, out of the plane of the diagram, to an extent dependent on its longitudinal velocity v_z (the direction of the classical current is opposed to that of the electron). The resulting motion of the electron in the x -direction, about the axis, will now cause a force to be exerted on it in the y -direction (by the field H_z), and this will be directed towards the axis. The radial acceleration produced by the

second half of the lens is in the opposite sense, so that the motion of the electron in the meridional plane is as depicted: the curvature of the path is a maximum at the centre of the lens, and decreases gradually until it leaves the field as a straight line. The point of intersection of this path with the axis gives the image point, I . The image will be rotated by an angle ϕ with respect to the inverted position which is normal in the case of such a converging system in light optics; ϕ is usually much less than 90° . When the magnetic field is in the reverse direction to that depicted, the electron will suffer a rotation of equal amount but in the opposite sense; its path in the meridional plane will be the same as that shown in Fig. 50 *a*. If the incident electrons are parallel to the axis, then in a perfect lens they will cross the axis in a point image the position of which will give the focal length of the lens.

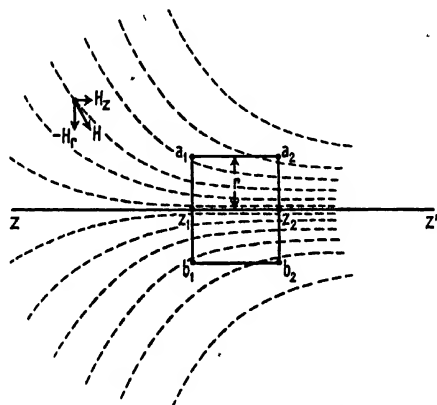


Fig. 51. Enlarged section of field in magnetic lens.

Approximate Expression for Focal Length. The first step towards finding expressions for the focal length and the rotation of the image is to determine the value of the radial field. Suppose Fig. 51 to represent a section of an inhomogeneous field of rotational symmetry in which $a_1 b_1$ and $a_2 b_2$ are two identical circular disks (of radius r) perpendicular to the axis of the field, $z z'$. They thus define a cylindrical volume in the field for which the Laplace condition must hold: as many lines must leave as enter it, in the absence of any source of flux within it. If the mean field strength (number of lines per unit area) across $a_1 b_1$ is H_1 , and that across $a_2 b_2$ is H_2 , then the net difference in the flux through the end disks is $\pi r^2 (H_2 - H_1)$, which must be equal to the number of lines entering the volume through its cylindrical sides. The area of the latter is $2\pi r (z_2 - z_1)$, and therefore the mean value of the radial field across them must be given by

$$H_r = \frac{-r(H_2 - H_1)}{2(z_2 - z_1)},$$

being negative since it is directed towards the axis, in the opposite sense to that in which r is measured. For an elementary unit of volume, a cylinder of vanishingly small length (dz), this may be written

$$H_r = -\frac{1}{2}r dH_z/dz. \quad (\text{IV.4})$$

It will be recognized that this is exactly analogous to the expression (III.12) found in the last chapter for the radial field in an electrostatic aperture lens by a more general argument; the two methods may be interchanged, since both depend on the application of Laplace's condition to the flux of lines of force, in the one case magnetic and in the other electrical. The second method is given here for the sake of the clearer physical picture it provides.

Discussing the motion of the electron in cylindrical coordinates (r , ϕ , and z), the radial field H_r will exert a rotational force on the electron on account of its axial velocity v_z , producing an acceleration $r\ddot{\phi}$, $\dot{\phi}$ being the angular velocity about the axis. We thus have

$$mr\ddot{\phi} = eH_r v_z = -\frac{1}{2}e r v_z dH_z/dz, \quad (\text{IV.5})$$

from IV.4. We may eliminate dz , since $v_z = dz/dt$, and obtain

$$\dot{\phi} dt = -(e/2m) dH_z.$$

Integrating from $-\infty$ to z , we have for the rotational effect of the field up to this point

$$\dot{\phi} = - \int_{-\infty}^z (e/2m) dH_z = (-e/2m) H_z. \quad (\text{IV.6})$$

This lateral motion of the electron is also normal to the z -component of the field, which therefore exerts a force on it in a direction at right angles to both, that is, radially. The radial force must be equal to the sum of the magnetic and the centrifugal forces, and so,

$$\begin{aligned} m\ddot{r} &= eH_z(r\dot{\phi}) + m(\dot{r})^2/r \\ &= \frac{-e^2 r H_z^2}{2m} + mr \frac{e^2}{4m^2} H_z^2 \\ &= \frac{-e^2 r}{4m} H_z^2. \end{aligned}$$

Thus the radial acceleration is

$$\ddot{r} = \frac{-e^2 r}{4m^2} H_z^2, \quad (\text{IV.7})$$

and, being negative, is directed towards the axis; because of the dependence on the square of H_z , this will be the case whichever the direction of the axial magnetic field in the first half of the lens. Since $\frac{1}{2}mv_z^2 = eV$, where V is the accelerating potential applied to the electron, this expression may also be written

$$\ddot{r} = \frac{d^2 r}{dt^2} = -\frac{er}{8mV} v_z^2 H_z^2, \quad \text{or} \quad \frac{d^2 r}{dz^2} = -\frac{e}{8mV} r H_z^2, \quad (\text{IV.8})$$

on elimination of t . In this form it is frequently referred to as the 'magnetic ray-equation', by analogy with the corresponding expression for the trajectory in an electrostatic field (II.16); unlike the latter, however, it contains the characteristic electron ratio, e/m .

Integrating IV.7, right through the lens, we have

$$\dot{r}_2 - \dot{r}_1 = - \int_{-\infty}^{\infty} \frac{e^2 r}{4m^2} H_z^2 dt = - \frac{e^2 r_m}{4m^2 v_z} \int_{-\infty}^{\infty} H_z^2 dz, \quad (\text{IV.9})$$

since $dz/dt = v_z$, the axial velocity of the electron, and we can assume that both this and r ($= r_m$) are constant across a short lens. This expression for the radial velocity is formally identical with that obtained for the electrostatic aperture lens (III.13), except that the factor $(e/2m)$ enters here in the second instead of the first power, as also does the axial field strength. In the same way equation IV.7 is equivalent to that obtained for the radial acceleration in both the aperture and the cylindrical lens (II.14).

Expression IV.9 shows that the deviation suffered by an electron is proportional to its distance from the axis, and hence the primary condition for paraxial image formation is fulfilled. As in ordinary Gaussian optics, we may now express the initial radial velocity in terms of the angle made with the axis by the incident beam

$$\dot{r}_1/v_z = dr/dz = \tan \theta_1 = r_m/p,$$

where r_m is the distance from the axis of the electron when passing through the lens, and p is the object distance (cf. Fig. 50). Similarly, writing q for the image distance, we have for the radial velocity on emergence from the lens

$$\dot{r}_2/v_z = -r_m/q,$$

directed towards the axis and therefore negative in sign. Substituting in IV.9 we then have

$$r_m/q + r_m/p = \frac{e^2 r_m}{4m^2 v_z^2} \int_{-\infty}^{\infty} H_z^2 dz,$$

and therefore
$$1/p + 1/q = \frac{e^2}{4m^2 v_z^2} \int_{-\infty}^{\infty} H_z^2 dz = 1/f, \quad (\text{IV.10})$$

on defining f , in the usual way, as the image distance when the object is infinitely distant. We thus have for the focal length of the short magnetic lens

$$f_0 = \frac{4m^2 v_z^2}{e^2 \int_{-\infty}^{\infty} H_z^2 dz}. \quad (\text{IV.11})$$

It follows that, as in light, the lens would have a different focal length for rays of different wave-length (\equiv velocity) if the magnetic field were constant. The advantage of the magnetic coil over the optical lens, how-

ever, lies in the possibility of varying its power continuously and over a wide range by control of the current in the windings. By this means the value of the field can be adjusted to maintain a constant focal length when different accelerating voltages ($\equiv v_z$) are employed. If, for practical convenience, we express the velocity of the electron in terms of the accelerating potential in volts, V (I.5), or of its radius of curvature (ρ) in a uniform magnetic field of strength H_a (I.1), the reduced expression is

$$f_0 = \frac{V}{0.022 \int_{-\infty}^{\infty} H_z^2 dz} = \frac{4(H_a \rho)^2}{\int_{-\infty}^{\infty} H_z^2 dz}. \quad (\text{IV.12})$$

In this form it resembles the approximate formula for the focal length of a symmetrical lens (III.30). As might be expected, f depends directly on the accelerating voltage applied to the electron beam and inversely on the square of the magnetic field strength. The faster the electrons the longer is the focal length, and the stronger the focusing field the smaller is f . For high accelerating potentials ($> 10,000$ volts) the relativistic correction must be applied (section 10); this leads to the corrected focal length $f = f_0(1 + (V_k/1000))$, where V_k is in kilovolts.

By direct integration of IV.6 we obtain corresponding expressions for the rotation of the image with respect to the object. Remembering that $v_z = dz/dt$, we have

$$\begin{aligned} \int_{-\infty}^{\infty} d\phi &= \phi = -e/2mv_z \int_{-\infty}^{\infty} H_z dz \\ \text{or,} \quad \phi &= 0.15/\sqrt{V} \int_{-\infty}^{\infty} H_z dz \\ &= -1/(2H_a \rho) \int_{-\infty}^{\infty} H_z dz. \end{aligned} \quad (\text{IV.13})$$

Thus the rotation of the image depends directly on the field strength and inversely on the axial velocity, and not on their squares as is the case with $(1/f)$. It will be noted also that the constants in the expression for the latter (IV.10) are the square of those in IV.13; but, because of the integration, the value of $(1/f)$ is not numerically equal to the square of the rotation, ϕ . The focal length is always positive, by virtue of its dependence on the square of the axial field, but the rotation will be positive or negative according to the direction of the latter. As in the case of the solenoidal field, the trajectory is not strictly reversible. An electron projected from I (Fig. 50) in the direction of those arriving from O will have a similar path in the meridian plane, but will suffer a rotation farther around the axis, as indicated by the broken line in the projection (a) on the $x-y$ plane.

The above treatment has assumed that the axial velocity of the electron, and its radial distance from the axis, are constant within the lens. Hence

the expressions for f and ϕ (IV.11, and IV.13) are valid only for paraxial conditions in a short lens, that is, where the object and image are far from the lens in comparison with the axial extension of the field. The strict treatment starts from a more precise expression for the flux $\psi(r, z)$ (or for the magnetic vector potential A) in terms of the axial field:

$$\psi(r, z) = \pi r^2 \left[H_z - \frac{r^2}{8} H_z'' + \dots \right],$$

which then gives the value of the field in the z -direction at any point off the axis in terms of the axial values:

$$H_{z(r,z)} = \frac{1}{2\pi r} \frac{\partial \psi(r, z)}{\partial r} = H_z - \frac{r^2}{4} H_z'' + \dots$$

This expression is equivalent to II.12 for the potential V_r in the electrostatic case, and the subsequent treatment follows the same general lines (cf. sections 18 and 26), giving an integral in which r cannot now be considered constant. Evaluation by successive approximation leads to a result for the focal length of a thick lens in the form of a series of terms, analogous to the expression III.23. The integral $e/8mV \int H_z^2 dz$ replaces that of $1/4\sqrt{V} \int V''/\sqrt{V} dz$; the accelerating potential V , being constant throughout the magnetic lens, appears outside the integral as in expressions IV.12 and IV.13. As the case of superimposed electric and magnetic fields is later fully treated by this method (cf. section 36), the detailed expression will not be discussed further here.

Experimental Verification. The validity of the result for the focal length of a short lens can be tested either by observation of the image distance for various object distances, and substitution in the ordinary optical formula

$$1/f = 1/p + 1/q,$$

or by numerical evaluation of f from IV.12. The first method was employed by Busch, who used a short magnetic coil to focus the emission from a cold cathode so as to form an image of an aperture on a fluorescent screen. The object-image distance $(p+q)$ was thus constant, and the coil was moved axially in order to vary p and q . He plotted the values of f thus obtained against the ratio $p/(p+q)$, with the result shown in Fig. 52; the agreement with the simple optical formula is excellent. He showed also that the

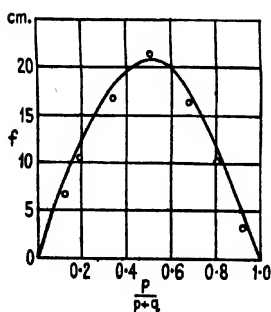


FIG. 52. Focal length of magnetic lens: comparison of experimental results (o) with theoretical curve. (Busch.)

focal length of the lens could be varied at will by varying the generating current I . This is to be expected from IV.12, since the field strength H is proportional to the current and the form of the axial distribution of

the field does not change with I , until saturation of the surrounding iron (if any) is approached. Busch gave the relation

$$1/f = kI^2$$

for an unshielded coil, where $k = 0.005$. As compared with electrostatic lenses, therefore, the current supply takes the place of the applied potential as the controlled variable for altering the power of the system. Johansson and Knecht confirmed Busch's result that the current-power curve for a magnetic lens was of parabolic form.

The latter authors also tested the accuracy of IV.12 by measuring the axial distribution of the field with a small search coil attached to a

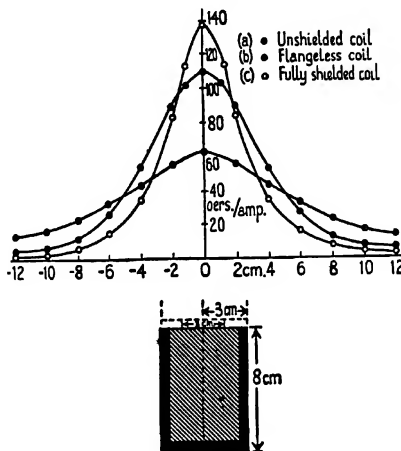


FIG. 53. Short magnetic lens: variation of field along axis, for three degrees of shielding.

ballistic galvanometer, and then carrying out a graphical integration to obtain a value of the integral $\int H_z^2 dz$. They found excellent agreement between the focal length thus calculated and that found by direct experiment. In this work the electron beam was accelerated by a potential difference of 750 volts. Tests at higher electron velocities have been made by Klemperer, using the beta-ray emission from thorium-B, which has a maximum at 150 kV., and a Geiger-Müller counter for recording the particles; and by the author, using beta-rays from radium-E and radon up to an equivalent voltage of 2×10^6 V., and a Wilson cloud-chamber for photographing the electron tracks directly (cf. section 64).

32. Form of Field in Short Magnetic Lens

The form of axial distribution of the field which occurs in a short lens depends on the degree of magnetic shielding of the generating coil. When a simple unshielded coil of approximately square cross-section of winding is employed, the field is found to extend well beyond the limiting planes of the coil, as shown in curve (a) of Fig. 53. On adding an iron sheathing

of a few millimetres thickness, as indicated by the heavy lines around the winding (shown in cross-hatching), the field is found to be concentrated more about the equatorial plane (curve (b)). The addition of ring shielding (dotted lines) to form a small magnetic gap at the centre causes a further contraction in width and increase in height of the field curve (c, Fig. 53). As the field strength is over a wide range directly proportional to the current, its value is usually quoted in gauss per ampere, as here. The distribution shown in the figure was obtained from a coil of 1,000 turns of copper wire wound with an inner diameter of 11 cm., an outer diameter of

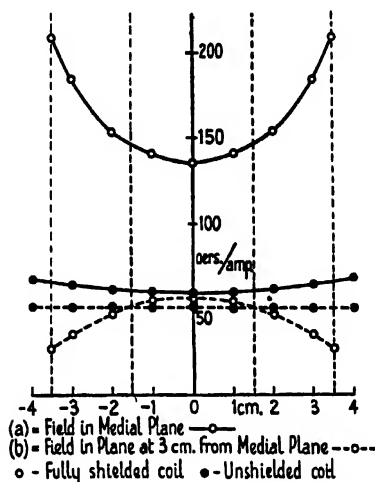


FIG. 54. Short magnetic lens: radial variation of field, with and without shielding.

26 cm., and an axial width of 5.2 cm.; but the general form is the same for coils of larger or smaller size. When large generating currents are used, the value of H_z in oersteds/amp. decreases as magnetic saturation is approached, and the peak of curve (c) is depressed.

Graphical integration of the distribution curves shows little difference in the value of $\int H_z dz$ for the three cases, and thus the rotation of the image is hardly changed by the addition of shielding. On the other hand, the focal length depends on the integral of the square of the axial field (IV.12), and we find that for curve (c) the value of $\int H_z^2 dz$ is nearly three times as great as for the unshielded coil. The advantage of shielding is clear; by using massive iron shielding in the construction of a beta-ray spectrometer system the author has obtained a concentration factor of 15 (section 64).

The radial distribution of the field also changes when iron shielding is applied to the coil. In Fig. 54 are shown the curves obtained on investigating the field in the mid-plane (a) and in the plane intersecting the axis at $z = 3$ cm., just outside the face of the coil (b), for both shielded and unshielded coils. The radial variation in the latter is small and positive with

increasing r ; only at greater axial distances than $z = 3$ does it take the opposite sign. On the other hand, the fully shielded coil shows a large positive radial variation in the mid-plane and a large negative change outside the coil. This behaviour is important in connexion with the production of spherical aberration.

The general form of the field is most clearly shown by a map of the magnetic equipotential lines, which are orthogonal to the lines of force (Fig. 55). The similarity to the distribution of potential in the symmetrical electrostatic lens will be seen at once on comparison with Fig. 41. The

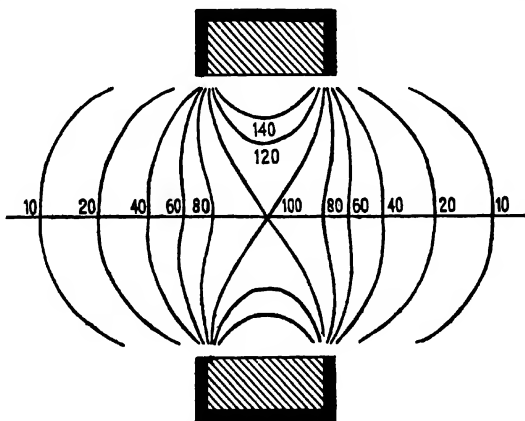


FIG. 55. Short magnetic lens: experimental map of field in meridian plane. (Klemperer and Miller.)

figures given against the lines in Fig. 55 indicate the magnetic potential with respect to that at the centre of the lens as 100 per cent.

The axial distribution of the field, and hence the power of the lens, is determined by the geometry of the coil, and attempts have been made to find an empirical relation between the number of ampere turns in the winding and the focal length, in order to avoid the labour of determining the field and making a graphical integration in each case. Ruska and Knoll, who were responsible for most of the pioneer work on shielded coils, introduced a coil factor F in the following manner. So long as saturation of the iron is not approached, we may write the field proportional to the number of ampere turns (IN), and from IV.12

$$(C \times IN)^2 = \int_{-\infty}^{\infty} H_z^2 dz = \frac{V}{0.022f},$$

where C is termed the coil constant. Thus

$$IN = \frac{1}{C} \sqrt{\frac{V}{0.022f}}. \quad (\text{IV.14})$$

In the case of a single turn of wire it is possible to derive directly the value

of the constant, which in this case is termed C_0 . If the turn is of radius r and carries a current I e.m.u., the value of the field at any point on its axis at a distance z from the centre of the coil is well known to be

$$H_z = \frac{2\pi r^2 I}{(r^2 + z^2)^{3/2}}.$$

Then, from the definition of C we can write

$$C_0^2 = \int_{-\infty}^{\infty} \frac{H^2}{(IN)^2} dz = 4\pi^2 \int_{-\infty}^{\infty} \frac{r^4}{(r^2 + z^2)^3} dz = \frac{3\pi^3}{2r}.$$

Substituting this value for C in IV.14, and expressing I in amps., we get for the required winding

$$IN = \sqrt{\frac{2rV}{3\pi^3 \times 0.022f}} = 220 \sqrt{\frac{V_k d}{f}}$$

on writing d , the diameter of the coil, in place of $2r$, and expressing V in kilovolts ($= V_k$). At high potentials the relativistic correction has to be applied, by multiplying the above expression by the factor $(1 + (V_k/1000))^{\frac{1}{2}}$.

In the case of a coil of square or rectangular cross-section, it is impossible to obtain an analytical expression for the field along the axis and hence for the constant C . We define instead a coil factor F_c , as the ratio of the coil constant C_0 for a single turn to the coil constant for the coil under discussion. We may then write

$$IN = 220 F_c \sqrt{\frac{V_k d}{f}}.$$

The value of F_c can then be determined by experiment for windings of different section. In the case of unshielded slab coils it has values of about 1.1 to 1.2; for shielded coils, where the presence of iron increases the field strength, its value is less than unity. From a knowledge of this empirical factor it is then possible to calculate the approximate focal length of any coil carrying a given current, or, alternatively, to design a coil to give any desired focal length.

Another method of approach is to find a simple function, that can be integrated, which fits the known axial potential distribution as accurately as possible, and then to derive an analytical solution of the expression for the focal length, IV.12. Glaser has attempted this for the type of magnetic lens made by Siemens for the electron microscope, in which the field distribution approximates closely to the form

$$H_z = \frac{H_0}{\{1 + (z/a)^2\}^\mu}, \quad (\text{IV.15})$$

where H_0 is the maximum value of the field in the lens, a is a parameter related to the half-width (d) of the H_z curve at half this maximum value,

$$a = \frac{d}{\mu/(\sqrt{2}-1)}, \quad (\text{IV.16})$$

and μ is a constant for a given type of coil (see Fig. 53). The value of μ

is about $\frac{1}{2}$ for an unshielded coil of the Siemens type and approaches 1 for shielded coils. On the other hand, Siegbahn has found a value of 3 for a different form of coil, when unshielded, tending to unity when shielded and approaching saturation.

The general equation, IV.15, does not yield to integration when substituted in IV.12. The case for which $\mu = \frac{1}{2}$ corresponds to the field distribution for a single circular turn or for a close wound coil of large diameter that approximates to the simple linear conductor. It has been fully treated by Wallauschek and Bergmann. Glaser took the more important case of $\mu = 1$ and showed that it leads to a simple expression for the focal length:

$$f = \frac{a}{\sin\{\pi/\sqrt{(1+k^2)}\}} \quad (\text{IV.17})$$

when measured from the principal plane, or

$$z_f = \frac{a}{\tan\{\pi/\sqrt{(1+k^2)}\}} \quad (\text{IV.18})$$

when measured from the mid-plane, as in earlier discussion above and corresponding to the mid-focal distance in the case of electrostatic lenses. The factor k is a parameter which characterizes the power of the lens and is given by

$$k^2 = \frac{eH_0^2 a^2}{8mV} = \left(\frac{H_0 a}{2H\rho}\right)^2. \quad (\text{IV.19})$$

These expressions then permit calculation of the spherical and chromatic aberration of the lens. They reduce to Busch's simple formula (IV.12) for small values of k , that is, for thin or weak lenses. Siegbahn has found good agreement with these formulæ in the case of a large coil. He has also treated the case of $\mu = \infty$, which he finds obtains in a semisolenoidal lens. In this case the expression for the axial field reduces to

$$H_z = H_0 e^{-(z/b)^2}, \quad (\text{IV.20})$$

where

$$b = a_\infty = d/\sqrt{(\log_e 2)}.$$

The solution proves to be of the same form as IV.17 and IV.18, with b replacing a , and k_∞ in place of k , if k_∞ is the value of IV.19 when a is replaced by b .

These are the only cases for which analytical solutions of the magnetic ray-equation have so far been found, apart from some simple distributions which have no practical importance, such as that of a rectangular H - z curve (see Wallauschek and Bergmann; also Preston).

33. The Magnetic Immersion Objective

As in the case of electrostatic lenses we understand by the term 'immersion objective' a system in which the object lies well within the magnetic field, which must therefore be of high focusing power. The special type of lens devised for this purpose may, of course, be used in the reverse sense of forming a very small image of a distant object, the image then falling

within the lens field. Special pole-pieces and coil windings are employed to obtain the high field strength required over a very short axial distance, but the focusing action and the electron paths are essentially the same as in the normal short magnetic lens. The approximate formulæ for focal length and image rotation previously derived will not apply in this case: although the axial velocity of the electron in the lens may still be taken as constant, the radial distance of the path now changes considerably and the electron may even cross the axis within the lens. The full expression for the focal length (corresponding to III.23 in the electrostatic lens) must

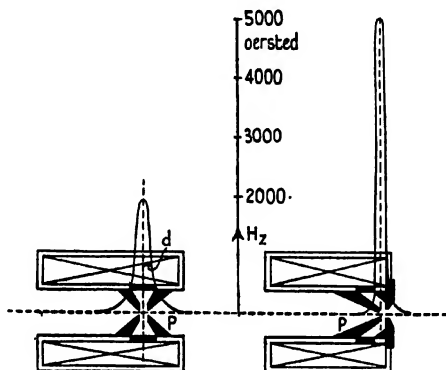


FIG. 56. Magnetic objective lens: construction and field distribution for two pole-forms. (Marton and Hutter.)

now be employed, unless the field approximates closely to the form for which Glaser has given an analytic solution (IV.15).

The short-focus objective was originally devised by Ruska for use in the magnetic electron microscope. It consists of a winding of considerable axial extension surrounded by a soft-iron shroud several millimetres in thickness. Water-cooling may be incorporated in order to allow of high generating currents, in which case the winding is sometimes divided into two sections. The iron shrouding is complete except for a small gap, into which pole-pieces are screwed (*P*, Fig. 56). These are made in the form of rings of semi-conical cross-section and may be fitted at or near the mid-plane of the lens, or at one end; in the developed form of the electron microscope the latter arrangement is preferred owing to the greater ease of insertion and adjustment of the object, which is mounted in a holder that sits in the upper pole-piece. The gap between the pole-pieces is normally filled by a brass spacing-ring (shown black), not only to obtain accurate separation of the poles but also to facilitate making the whole system vacuum-tight. The best soft iron must be used as the field is now concentrated into the gap between the poles, which may be as small as 2 or 3 mm., and saturation of the iron must be avoided as far as possible.

Measurement of the field distribution within the lens is now a matter of some difficulty, owing to the small dimensions of the poles, and reliable

measurements cannot be made on a large-scale model on account of the importance of saturation effects in the iron. A number of devices have been described for overcoming these limitations, since detailed investigation of the form of the field is essential to the design of lenses with shorter focal length and less aberration. Klemperer proposed a mechanical arrangement for oscillating a small search coil in the field, amplifying the output in a resonant amplifier. Other workers have relied upon rotating a coil at high speed at the end of a shaft driven by a synchronous motor. Marton has refined the normal method by constructing a search coil of less than 1 mm.^3 in volume, which could readily be inserted between the pole-pieces; the effect produced in it by reversing the current in the lens was recorded with a ballistic galvanometer. Carrying this method further, Dosse has made a coil of less than 0.05 mm.^3 in volume, having 100 turns of fine wire, with which he claims an error of less than 0.2 per cent. in measurement of the field.

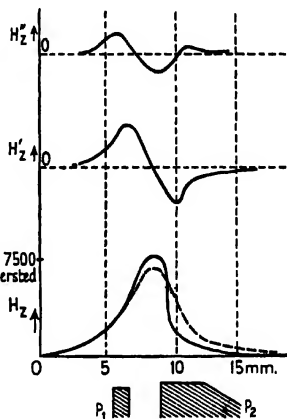


FIG. 57. Magnetic objective lens: axial variation of field and its first and second derivatives for special pole-forms. (Dosse.)

The form of the axial field strength found by Marton is shown in Fig. 56, for two arrangements of the pole-pieces. The greater concentration obtained with pole-pieces fixed at one end of the coil results in a very short focal length, since f depends inversely on $\int H_z^2 dz$, and this type is employed in the R.C.A. electron microscope. Dosse measured the field between two unsymmetrical pole-pieces, of the shape shown above, Fig. 57 (P_1, P_2). The peak of the curve is not so sharp as in Marton's lenses and has less symmetry. The form of the curves for the first and second derivatives of H_z are also given: Dosse claims that his measurements are accurate enough to warrant direct graphical evaluation, although the H'_z curve has to be smoothed from a considerable scatter of the experimental points. On doubling the value of the current in the lens, he found that the peak of the H_z curve was depressed, owing to saturation effects, as indicated by the broken line; the form of the field is now more symmetrical.

The approximate formula (IV.11) for the focal length of the short magnetic lens does not give a reliable result for the immersion type of lens, even when the object is situated outside the field. The power of the lens is now so great that the simplifying assumptions made in section 31 are no longer even approximately valid; in particular, the radial distance of the electron from the axis is by no means constant within the lens. In very strong lenses the trajectory may cross the axis within the field, and recross it again later. The actual focal length may differ by as much as a factor of 3 from that given by equation IV.11.

In order to obtain a more accurate expression for f in the case of the immersion objective, Glaser has produced an analytical solution of the magnetic ray equation (IV.8) by finding a function that fits closely to the known axial field distribution and can be integrated. This is a simple case of the more general expression (IV.15) given in the previous section, with the value of unity assigned to the constant μ ; we then have

$$H_z = \frac{H_0}{1 + (z^2/a^2)}, \quad (\text{IV.21})$$

and, by IV.16, the parameter a now takes the same value as d , the half-value half-width of the field distribution curve (see Fig. 60). The general solution for the electron trajectory is found to be:

$$r_z = a \sqrt{1 + \left(\frac{z}{a}\right)^2} \left[C_1 \sin \left\{ \sqrt{1 + k^2} \operatorname{arccot} \left(\frac{z}{a} \right) \right\} + C_2 \cos \left\{ \sqrt{1 + k^2} \operatorname{arccot} \left(\frac{z}{a} \right) \right\} \right], \quad (\text{IV.22})$$

where C_1 and C_2 are two arbitrary constants of integration which can be so determined that the path satisfies any desired boundary conditions. Giving them values corresponding to an incident beam parallel to the axis, the focal point is located from the general expressions

$$z_f = \frac{a}{\tan \{ n\pi / \sqrt{1 + k^2} \}}, \quad (\text{IV.23})$$

$$f = \frac{a}{\sin \{ n\pi / \sqrt{1 + k^2} \}}, \quad (\text{IV.24})$$

where z_f is measured from the mid-plane and f from the principal plane of the lens. The parameter k has the value given by IV.19, and n may be any positive integer. Single images are formed by the lens when $0 < k^2 < 3$ and $n = 1$, and the above expressions then reduce to those given earlier (IV.17 and IV.18). It follows that the minimum value of f for this form of field will occur when $k^2 = 3$ and the value of the sine term is unity. Then $f = a$, and $z_f = 0$: the focal point will be in the centre of the lens, and the principal plane at the position where H_z has the value of $H_0/2$. For higher values of k^2 the rays will cross the axis more than once within the lens, forming intermediate images. This type of field distribution is closely obeyed by all lenses of the immersion objective type, in which the pole-pieces are of small internal diameter and are closely spaced, so that the field is concentrated into a very small volume. It is possible, therefore, to use the above expressions with some confidence in theoretical discussion of the aberrations of such lenses, as has recently been done by both Dosse and Glaser (cf. section 39).

The position of the other cardinal points can also be deduced on the same assumptions, and this is especially valuable because their experimental location is a matter of difficulty. Ruska investigated the problem with the arrangement indicated in Fig. 58 a, where O is a wire-gauze object supported on a ring in the path of a high-voltage electron beam, above the

pole-pieces P of the lens (not otherwise shown). The object distance p could be varied by moving the supporting ring along the axis, and for each distance he determined the lens current ($\equiv \int H_z^2 dz$) required to produce a sharp image on the screen S , at a fixed image distance q . The results were expressed in the form of curves for the magnification (M_0) and lens current (I_0) against object distance, for various experimental conditions. Typical curves obtained for 60 kV. electrons are shown in Fig. 58 *b* (full lines); both

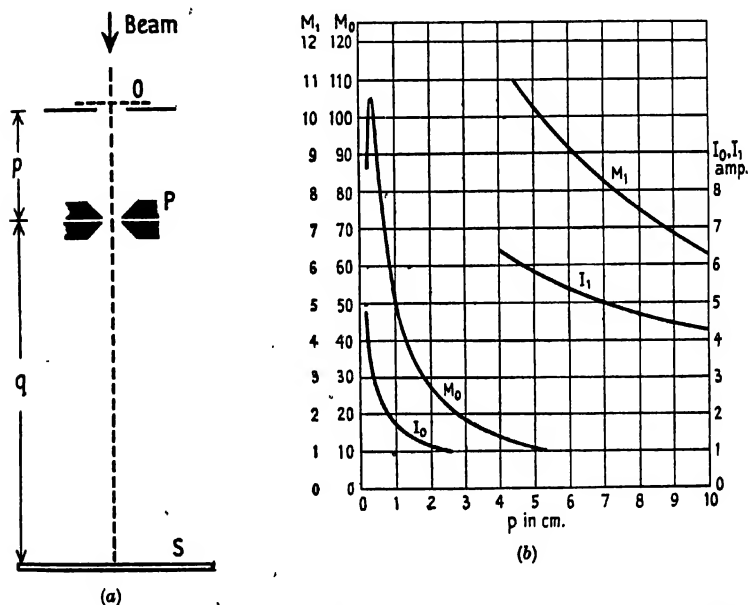


FIG. 58. Magnetic objective lens: (a) experimental arrangement. (b) magnification (M) and current (I) for varying object distance (p) and a 60 kV. beam. (Ruska.)

M and I : iron-clad coil.

M_1 and I_1 : unshielded coil.

M_1 on ten times the scale of M .

rise rapidly for object distances of less than 2 cm. The magnification, however, drops sharply at a value of $p = 3$ mm. on account of saturation of the pole-pieces. The broken lines show the values for the magnification M_1 (on ten times the vertical scale) and lens current I_1 (original scale) for the exciting coil without iron shrouding. These results were obtained with pole-pieces of 5 mm. diameter; the power of the lens increased with diminishing diameter (D), and Fig. 59 shows the form of the relation. A minimum focal length of 3 mm. was obtained with this type of lens.

Ruska attempted to determine the position of the principal planes from his measurements of object distance and magnification, by the normal light optical methods. He found that no consistent conclusions could be

obtained within the limits of his error in measurement (2 per cent.), indicating that the principal planes lay very close together. The fact that the observed focal length was usually greater (by a millimetre or less) than that calculated for the equivalent thin lens indicated that the principal

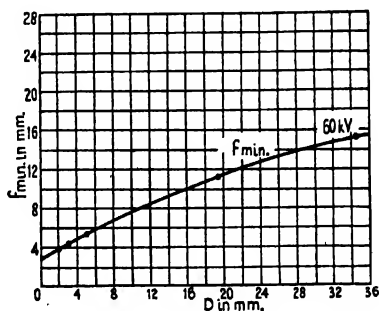


FIG. 59. Magnetic objective lens: variation in minimum focal length (f_{\min}) with pole-piece diameter (D) for a 60 kV. beam. (Ruska.)

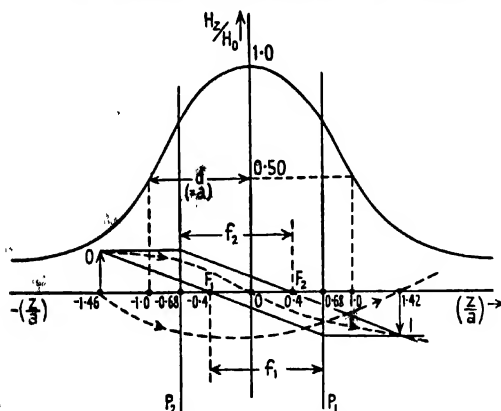


FIG. 60. Magnetic objective lens: form of field, ray paths (dotted), and principal planes,

$$\text{when } H = \frac{H_0}{1 + (z^2/a^2)}. \quad (\text{Glaser.})$$

planes were crossed over near the centre of the lens. These results were obtained for the unshielded lens coil; for the lens with pole-pieces it proved impossible to draw any conclusions at all, owing to the object falling within the field when such short focal lengths were employed. A satisfactory experimental method for strong objectives has not yet been devised.

Glaser, however, has determined the position of the cardinal points by calculation from expressions IV.23 and IV.24, for the special type of field assumed in their derivation. The form of these expressions shows at once that $f > z_p$ for all values of k , and hence that the principal planes are indeed crossed over. The position of the four main cardinal points, for a

value of $k^2 = 1.6$, is shown in Fig. 60. For a focal length of 5 mm. the principal planes prove to be separated by a distance of 6.8 mm. The general expression for their separation is shown by Glaser to be

$$2z_p = 2a \cot \frac{\pi}{2\sqrt{(k^2+1)}}. \quad (\text{IV.25})$$

Hence in strong lenses ($k > 1$) the separation increases, in terms of a the field-form parameter, and for weak lenses ($k < 1$) it tends to zero. The full lines in Fig. 60 show constructional rays; the paths actually followed by electrons are curved in this plane, as indicated by the broken lines. Glaser was able to show that the normal optical formulae and constructional methods were none the less valid in these special conditions, when the object or image is deeply immersed in the field.

34. Ray-tracing in Magnetic Lenses

Considerably less attention has been paid to methods of ray-tracing in magnetic lenses than to the corresponding electrostatic problems. The rotation of the beam about the axis provides an additional complication in the magnetic field, and in any case it is not possible to determine for it a set of equipotential surfaces with the ease and directness that electrostatic fields allow. One method of approach, indeed, is to convert the magnetic lens into its electrostatic counterpart, but this is by no means a simple process, as will be made clear later. It has been more usual to attempt to adapt the methods used in the electrical case, either directly or with the aid of the concept of magnetic vector potential. Owing to the different properties of this quantity as compared with the electrostatic scalar potential (see next section), it is not allowable to use identical methods in the two cases.

The most direct and rapid procedure is to use the paraxial ray equation for the magnetic field (IV.8) in conjunction with an experimental determination of the distribution of the field along the axis, H_{0z} . Siday has calculated trajectories through a thick lens, formed by a short solenoid, by making a step-by-step integration of the equation from such a distribution curve. A less laborious method has been used by the author, akin to that given for electrostatic lenses in section 18; it also has an advantage in accuracy, as it derives the values of the field parallel to the axis (H_{rz}) at each point of the path, instead of assuming it equal to that at the corresponding point on the axis (H_{0z}).

The form of the ray equation used is that of IV.8:

$$r'' = -(e/8mV_0) H_{rz}^2 r = (0.022/V_0) H_{rz}^2 r, \quad (\text{IV.26})$$

where V_0 is the accelerating potential applied to the electrons (in volts). Expressing the field in the r - and z -directions at the point (r, z) in terms of the value at the axis, as given by the series (IV.45 and IV.46) derived in the next section, a more exact treatment yields the equation

$$r'' = -(e/8mV_0) \left(H_{0z}^2 - \frac{r^2}{2} H_{0z}'' H_{0z}' + \frac{r^4}{16} H_{0z}''^2 \right) r, \quad (\text{IV.27})$$

neglecting higher orders of derivatives of H . The value of H_{0z} is found by an experimental determination of the field along the axis, at 1-cm. intervals; the values of H_{0z}' and H_{0z}'' are found from second and fourth differences as in the electrostatic method. Assuming an incident beam of given direction, so that r_0 and r_0' are known, the value of the radial acceleration r_0'' imposed on an electron entering the field can be calculated from IV.27. The value of the radial distance and radial velocity (r_1 and r_1') at the end of this segment of the lens is then found from the relations

$$r_1 = r_0 + r_0' \Delta z + \frac{1}{2} r_0'' \Delta z^2, \quad \text{and} \quad r_1' = r_0' + r_0'' \Delta z, \quad (\text{IV.28})$$

where Δz is the segment length chosen. These values, with the next value of H_{0z} , are then used to find the value of r_1' ; and so on, segment by segment, through the lens. The number of the segments, and hence their axial length, are chosen according to the degree of accuracy required; as in the electrostatic case, 15 to 20 points are usually sufficient. If the number of segments is made greater than 20, then differences of higher order should be introduced into equations IV.28 so that the accuracy of r and r' is increased accordingly; the necessary expressions were given for the electrostatic case (II.17, II.18). By taking the incident ray parallel to the axis from each side of the lens in turn, the position of the principal foci and hence of the principal planes can be found at once. However, since these latter are close together, it is advisable to plot at least one other pair of rays through the system as a check on their location. When a rapid indication of the focal length only is required, it is sufficient to assume that ${}_z H_{rz} = H_{0z}$ and to employ the simple expression IV.26. A comparison of the results obtained by the rough and the more accurate procedure is given in Table II below, together with experimental values.

The path calculated by this method is that followed by an electron in a meridional plane rotating about the axis with it. The rotation can be found by a similar step-by-step procedure, from the equation for the angular velocity:

$$\dot{\phi} = -(e/2m) {}_z H_{rz}, \quad (\text{IV.29})$$

but it is usually not a matter of interest. The method also assumes that the velocity of an electron parallel to the axis is constant through the lens, as determined by the accelerating potential V_0 . It is therefore less accurate when applied to lenses of high power. It would be possible, but laborious, to calculate the correction to be applied to the velocity at each point of the path, from the expressions for the tangential velocity (IV.29) and the radial velocity (IV.28), and thus make the method better applicable to immersion objectives.

Goddard and Klemperer have recently described a more exact method, which is based on a similar mathematical procedure. They prefer, however, to integrate with respect to time and not with respect to z as in the method just described. For this purpose they take as starting-point the three equations of motion of the electron, in terms of the magnetic vector potential A (see next section) and its derivatives with respect to r and z .

The value of A for each point of the trajectory is found by numerical evaluation from an experimental map of the field strength H_z in the median plane. For a ray of given initial direction, r and \dot{r} , z and \dot{z} , the successive values of the radial and axial accelerations are found with the aid of formulae derived again from the calculus of finite differences, but of greater accuracy than II.17 and IV.28:

$$r_{n+1} \approx r_n + r_{n-2} - r_{n-3} + \frac{\omega^2}{4} (5\ddot{r}_n + 2\ddot{r}_{n-1} + 5\ddot{r}_{n-2}),$$

$$\text{and } r_{n+1} \approx r_n + r_{n-4} - r_{n-5} + \frac{\omega^2}{48} [67\ddot{r}_n - 8\ddot{r}_{n-1} + 122\ddot{r}_{n-2} - 8\ddot{r}_{n-3} + 67\ddot{r}_{n-4}],$$

the second and more exact expression being used where the field gradient is steep and the path correspondingly of sharp curvature. The time unit ω is chosen to accord with the degree of accuracy required in the trajectory, and the subscripts $n, n-1, \dots$ refer to the values of r or \dot{r} at previous time intervals. By this means the conditions along a section of the path and not simply at the point considered, are taken into account. A similar relation applies for determining z . It is then possible at once to find the values of $r, \dot{r}, \ddot{r}, z, \dot{z}, \ddot{z}$ at each successive unit of time and hence to obtain the whole trajectory. The angular velocity $\dot{\phi}$ and angle of rotation ϕ can also be found from the third equation of motion. The expressions are in such a form that they can be rapidly evaluated with an electrical calculating machine.

The method has the advantage that it can be employed on rays starting with a component of velocity about the axis (skew rays) and for those of large aperture. In addition, the accuracy may be checked at each stage of the procedure: the radial, axial, and rotational velocities may be rapidly integrated numerically up to any point, and the sum of their squares evaluated ($\dot{z}^2 + \dot{r}^2 + \dot{\phi}^2$). This sum should be constant throughout the path, being directly proportional to the energy of the electron, which is unchanged by the action of the magnetic field. On account of this control, therefore, the method is the most reliable yet devised for investigating the errors of lenses by ray-tracing.

In order to test the degree of accuracy obtained Goddard and Klemperer investigated by direct experiment the electron paths in a simple magnetic lens, using an adaptation of the apparatus described in section 23 for investigating electrostatic lenses. Electron beams of varying inclination to the axis were projected into the lens from a special electron gun, and their subsequent path was found by observing the spot formed on a fluorescent screen which could be shifted to and fro along the axis. By this means it was possible to follow the beam right through the lens, since the presence of the screen in the field does not disturb it as it does an electrostatic field. The rotation ϕ of the beam was observed as well as its final inclination to the axis (and hence the focal length) for beams of different initial inclination, $\tan \theta$. In Table II the results are shown in comparison with the values computed by the method described above;

the mid-focal length MF is also compared with the value obtained from the approximate formula (IV.11), which is equivalent to setting $H_{rs} = H_{0s}$ in IV.26. The agreement between experiment and ray-tracing is good, both for the rotation and the focal length (see also Fig. 61). On the other hand, the Busch formula is already significantly in error for the beam of $\tan \theta = 0.058$, or only 3° divergence from the axis, in spite of the fact

TABLE II

Comparison of Experimental and Computed Electron Trajectories in an Unshielded Short Magnetic Lens at Various Apertures ($\tan \theta$)

(G. & K. = Goddard and Klemperer's method; B. = Busch's paraxial formula (IV.11); C. = author's method (IV.27).)

$\tan \theta$	Angular rotation (ϕ)		Mid-focal length (z_f) (cm.)			
	Experiment	Computation (G. & K.)	Experiment	Computation		
				G. & K.	B.	C.
0.034	59°	60°	27.0	23.8	26.8	25.3
0.068	66°	61°	23.2	22.2	26.4	23.1
0.101	67°	63°	18.0	18.4	26.5	19.1
0.125	67°	65°	13.3	15.3	26.4	15.9

that the lens is by no means of high power: the maximum field strength on the axis (H_0) was 16.5 oersteds and the incident electrons of 500 e.V. energy. The final column gives the values computed by the author's method (IV.27). They differ little from the other computed values, but agree rather better with experiment for paraxial rays. For slightly different conditions the true focal length, f , as well as z_f , was calculated from the expressions given by Glaser (IV.23 and IV.24), assuming the field to have the form for which these were evaluated (IV.21), and comparison was then made with experiment and the new method of computation. For both f and z_f , the latter gave (paraxial) values slightly greater than those of Glaser and 3 per cent. below those found by experiment. In agreement with other work the principal planes proved to be crossed over, at a separation of 1.0 cm., when the focal length was 10.9 cm. (cf. section 33). The circumstance that computation usually gave values for f smaller than those found by experiment was ascribed to a uniform error in measurement of the field, due to the finite size of the search coil.

This method of computation, therefore, gives results of a degree of accuracy not previously obtained, and holds out possibilities for investigating the aberrations of magnetic lenses by ray-tracing, as in light optics. Goddard and Klemperer traced a number of parallel rays at different radial distances in order to determine the spherical aberration, and compared the results with the experimentally determined focal lengths (Fig. 61). The aberration was found to be positive, giving shorter focal lengths for marginal than for paraxial rays (cf. section 39); the computed

values were 3 per cent. in error at the axis and 9 per cent. too low for a radial distance of 3 cm., equal to about half the radius of the lens.

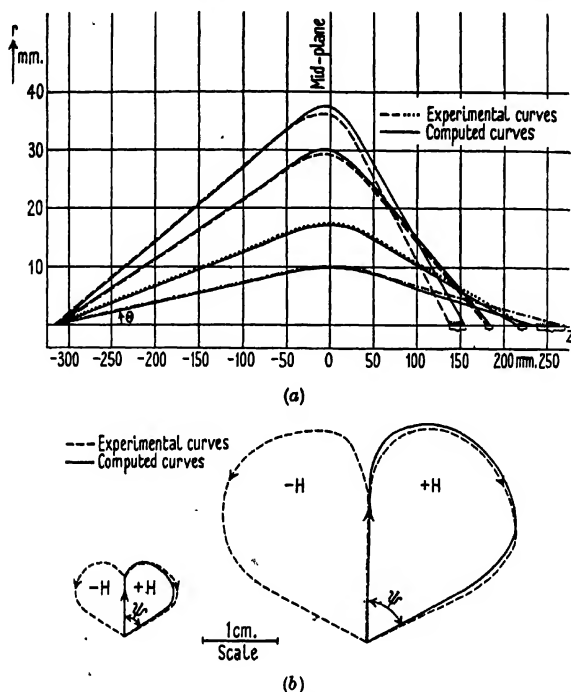


FIG. 61. Electron trajectories in unshielded magnetic lens at various apertures: (a) in meridional plane (b) in projection on transverse plane. (Goddard and Klemperer.)

35. Magnetic Vector Potential; Equivalent Electrostatic Lens

We have so far specified the magnetic field by reference to the field strength, or flux per unit area, H . An alternative approach employs the concept of magnetic vector potential, A , which is related to H in the same way as H is related to the generating current density i . In elementary discussion it is more convenient to use the physically more intelligible quantities, the flux and field strength. We shall now indicate the relations between A and H in the case of fields of rotational symmetry, and the ways in which the vector potential is especially helpful in lens problems.

The curl function of a field may be defined in terms of the work done in carrying unit pole around the current-carrying wire that produces the field. Irrespective of the path followed, the line integral of the field (work done) taken around a closed curve is given by

$$\oint H ds = 4\pi I, \quad (\text{IV.30})$$

if I is the current in the wire, as was shown by Oersted. The quantity,

$\text{curl } H$, is then defined as the line integral per unit area, and is obtained on dividing IV.30 by the area (σ) enclosed by the path:

$$\text{curl } H = \int H \, ds / \sigma = 4\pi I / \sigma = 4\pi i, \quad (\text{IV.31})$$

where i is now the mean current density enclosed by the path. When the path is not linked by any current, we have the corresponding expression to Laplace's equation of the electrostatic field:

$$\text{curl } H = 0.$$

The vector potential is formally defined in the same way with respect to H :

$$\text{curl } A = H \quad (\text{IV.32})$$

or

$$\int A \, ds / \sigma = H. \quad (\text{IV.33})$$

In general the closed curve along which the line integral is taken will not be confined to one of the planes of reference. The components of A can then be found, in normal rectangular coordinates, by projecting the curve on to the respective planes. In the present case it is convenient to work in cylindrical coordinates and, if no current links the curve, the components of the curl of A and of the curl of H are then

$$\left. \begin{aligned} \text{curl } H_z &= \frac{1}{r} \left(\frac{\partial(rH_\phi)}{\partial r} - \frac{\partial H_r}{\partial \phi} \right) = 0; \\ \text{curl } H_r &= \frac{1}{r} \frac{\partial H_z}{\partial \phi} - \frac{\partial H_\phi}{\partial z} = 0; \\ \text{curl } H_\phi &= \frac{\partial H_r}{\partial z} - \frac{\partial H_z}{\partial r} = 0; \end{aligned} \right\} \quad (\text{IV.34})$$

and

$$\left. \begin{aligned} H_z &= \text{curl } A_z = \frac{1}{r} \left(\frac{\partial(rA_\phi)}{\partial r} - \frac{\partial A_r}{\partial \phi} \right); \\ H_r &= \text{curl } A_r = \frac{1}{r} \frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z}; \\ H_\phi &= \text{curl } A_\phi = \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r}. \end{aligned} \right\} \quad (\text{IV.35})$$

In magnetic lenses the field is usually of rotational symmetry, in which case all derivatives of H with respect to ϕ vanish, the components A_z and A_r are zero, and we have

$$\left. \begin{aligned} H_z &= \frac{1}{r} \frac{\partial(rA_\phi)}{\partial r} \\ H_r &= -\frac{\partial A_\phi}{\partial z} \end{aligned} \right\} \quad (\text{IV.36})$$

Substituting in the last equation of IV.34, we obtain

$$\frac{\partial^2 A_\phi}{\partial r^2} + \frac{1}{r} \frac{\partial A_\phi}{\partial r} - \frac{A_\phi^2}{r^2} + \frac{\partial^2 A_\phi}{\partial z^2} = 0, \quad (\text{IV.37})$$

which corresponds to Laplace's equation (II.3) for an electrostatic field of rotational symmetry; it must be emphasized, however, that the two conditions are not identical. From this we may proceed, as indicated in the next section, to obtain an expression for the vector potential at any point in terms of its value on the axis, analogous to II.12 for the electrostatic potential.

From the definition of A it follows that, for a finite surface,

$$\int H_z d\sigma = \int \text{curl } A d\sigma = \int A ds,$$

that is, the surface integral of the normal component of the field over a given surface is equal to the line integral of the vector potential taken round the closed boundary of this surface (Stokes's theorem). In general it is not possible to measure A directly nor to obtain a simple expression relating it to other parameters of the field, but in the special case of rotational symmetry here considered it can be related at once to the flux. We suppose the surface to be a disk of radius r_0 set normal to the axial component of the field H_z ; its boundary is then of length $2\pi r_0$, and encloses an area of πr_0^2 . Hence

$$\oint A ds = 2\pi r_0 A = \int_0^{r_0} 2\pi r H_z dr = \psi, \quad (\text{IV.38})$$

where ψ is the total flux through the area, obtained by integrating the product of field strength and area from the centre ($r = 0$) to the circumference ($r = r_0$) of the disk. We may thus write

$$A = \psi/2\pi r_0 = (1/r_0) \int_0^{r_0} H_z r dr. \quad (\text{IV.39})$$

In magnetic lenses, therefore, the magnetic vector potential in a given region is given by the total normal flux through a small area enclosing it, divided by the perimeter of this area. In the special case when the field is uniform this reduces to

$$A = \pi r^2 H_z / 2\pi r = (r/2) H_z.$$

The equipotentials of A are then concentric cylinders around the axis; at the axis, $A = 0$. When the axial distribution of the field has been measured with a search coil, the distribution of vector potential may be found by graphical integration from IV.39. In order to avoid this extra step, however, methods directly involving H are in practice used in preference to vector potential methods whenever possible.

It will be shown in the next section that the motion of an electron in combined electric and magnetic fields can be described in terms of a 'meridional potential' (Q), distributed in the meridional plane which is

considered to rotate about the axis with the angular velocity ($\dot{\phi}$) of the electron. Once Q is known, the path of the electron can be traced through successive stages of refraction at the equipotential lines in this plane. If V is the accelerating potential applied to the electron, then

$$Q_{rz} = (e/m)V - r^2\dot{\phi}^2/2 = (e/m)V - \frac{1}{2}\{(e/m)A + C/r\}^2, \quad (\text{IV.40})$$

where the quantity C is given by

$$C = r^2\dot{\phi} - (e/m)rA,$$

and is determined by the initial rotation of the electron about the axis (cf. footnote, p. 107). For rays starting from a point on the axis C has the value zero, and we may write

$$A = (m/e)\sqrt{[2\{(e/m)V - Q\}]}, \quad (\text{IV.41})$$

giving an expression for the magnetic vector potential. This shows that equipotentials of A are also equipotentials of Q ; they have the general shape of the lines or surfaces of equal H , but are more closely spaced. Once values of A have been determined, the corresponding equipotentials of Q can be mapped and the trajectory of the electron traced (or calculated) from point to point, using Snell's law for the refraction at each boundary. Alternatively, Dosse has used a method comparable with that used in automatic ray-tracing in electrostatic fields: from the value of the normal component of the gradient of Q the force on the electron is found and a (parabolic) segment of path calculated until it reaches the next equipotential, consecutive segments being joined tangentially at each boundary in order to obtain the complete trajectory. This is probably as accurate as the method of Goddard and Klemperer but far more laborious.

The same avenue of approach also gives the form of that electrostatic lens which would have the same effect as a given magnetic field. Indeed, it is found at once by dividing the meridional potential Q by (e/m) . Equation IV.41 thus gives the equivalent electrostatic potential U :

$$U = V - (e/2m)A^2. \quad (\text{IV.42})$$

As V is constant in the normal type of magnetic lens, it follows that the equipotential surfaces of U will also be equipotentials of A . At the axis $A = 0$, and its value increases with distance radially, so that the value of U must decrease in this direction. That is to say, the field behaves as if the refractive index decreased with the radius; the lens may be considered as made up of hyperboloidal shells of decreasing index (or 'optical density') arranged around the axis. Figure 62 shows the form of the equipotentials in the equivalent electrostatic lens and, by shading, the diminution of refractive index away from the axis; near the axis the decrease in power is roughly proportional to the square of the distance. The electron, being a negative particle, will be deviated towards the axis in what is relatively a potential trough; thus, as already stated, the magnetic lens is always a converging system. By this method any magnetic lens may be trans-

formed into its electrostatic equivalent, in which trajectories may then be traced by one of the procedures described in Chapter II.

It is clear that this type of system is very different in the general form of its field from the electrostatic lenses considered there and in Chapter III. If anything, it is even more remote from any glass optical focusing arrangement. As Gabor has pointed out, however, an equivalent optical model could be made from transparent hyperboloidal shells. When rotated at high speed it would have exactly the same effect on light rays as the magnetic lens has on electrons: the light would travel in spiral paths

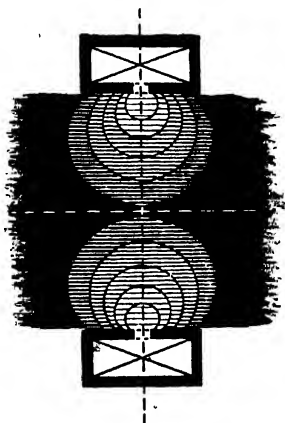


FIG. 62. Equivalent refractive index in magnetic lens (after Gabor).

(Fizeau effect). In practice, the speeds required to produce such a convection of light in a moving medium would be so high that the material would break up by centrifugal force.

*36. *Superimposed Magnetic and Electrostatic Fields: General Case*

Having considered as separate cases the electrostatic and the magnetic lens system, we now treat the general case in which electrons are under the simultaneous influence of both types of field. As stated in section 5, the total force on an electron is then given by adding vectorially the electric and magnetic forces. The electrostatic case was fully treated in sections 18 and 28, and it remains to obtain a general expression for the combined fields.

Commencing from a special solution of Laplace's equation for electrostatic fields of rotational symmetry,

$$V_{rs} = (1/2\pi) \int_0^{2\pi} f(z + i \sin \alpha) d\alpha,$$

P

it was shown in section 18 that the potential V_{rz} at any point (r, z) could be expressed in terms of the axial potential and its derivatives:

$$V_{rz} = V_z - \frac{r^2}{2^2} V_z'' + \frac{r^4}{2^2 \cdot 4^2} V_z^{IV} - \dots (-1)^n \frac{1}{(n!)^2} \left(\frac{r}{2}\right)^{2n} V_z^{(2n)}, \quad (\text{IV.43})$$

where primes indicate derivatives with respect to z , as before. The equivalent approach for the magnetic field begins with expression IV.37, in terms of the vector potential A ,

$$\frac{\partial^2 A}{\partial r^2} + \frac{1}{r} \frac{\partial A}{\partial r} - \frac{A^2}{r^2} + \frac{\partial^2 A}{\partial z^2} = 0.$$

The solution for fields of rotational symmetry is

$$A = (i/2\pi) \int_0^{2\pi} \Phi(z + ir \sin \alpha) \sin \alpha \, d\alpha,$$

where Φ is the scalar magnetic potential. Proceeding as before, this yields an expression for A in terms of the axial values

$$A = -\frac{r}{2} \Phi' + \frac{r^3}{2^3} \frac{\Phi'''}{2!} - \frac{r^5}{2^5} \frac{\Phi^{(5)}}{2! 3!} + \dots \frac{(-1)^n}{n!(n-1)!} \Phi^{(2n-1)} \left(\frac{r}{2}\right)^{2n-1};$$

$$\text{or } A = \frac{r}{2} H_{0z} - \frac{r^3}{16} H_{0z}'' + \frac{r^5}{6 \times 64} H_{0z}^{IV} - \dots \frac{(-1)^n}{n!(n+1)!} H_{0z}^{(2n)} \left(\frac{r}{2}\right)^{2n+1}, \quad (\text{IV.44})$$

since $H_z = -\Phi'$. From the relation between A and H_z previously given (IV.36), we obtain on differentiation with respect to r the corresponding and more useful expression for ${}_r H_{rz}$ in terms of the axial field H_{0z} :

$${}_r H_{rz} = H_{0z} - \frac{r^2}{4} H_{0z}'' + \frac{r^4}{64} H_{0z}^{IV} - \dots \frac{(-1)^n}{(n!)^2} \left(\frac{r}{2}\right)^{2n} H_{0z}^{(2n)}. \quad (\text{IV.45})$$

Similarly we get for the radial field ${}_r H_{rz}$ the series

$${}_r H_{rz} = -\frac{r}{2} H_{0z}' + \frac{r^3}{16} H_{0z}'' - \frac{r^5}{384} H_{0z}^{IV} + \dots \frac{(-1)^n}{n!(n-1)!} H_{0z}^{(2n-1)} \left(\frac{r}{2}\right)^{2n-1}. \quad (\text{IV.46})$$

It will be seen that the series for ${}_r H_{rz}$ is of the same form as that for the electrostatic case (IV.43). It is to be noted, however, that the magnetic field strength is a vector quantity, so that the total field at the point (r, z) is given by the sum of the vectors ${}_r H_{rz}$ and ${}_z H_{rz}$; and since the former increases with the radius more rapidly than the latter decreases, the total field must increase towards the perimeter of the lens, as is found experimentally. It is clear also from the last two equations that the simple relation derived in section 31,

$$H_r = -(r/2) dH_z/dz,$$

is valid only so far as the first term of the series is concerned. The present treatment does not involve the radial field, but a similar approximation is made in regard to the series for ${}_r H_{rz}$.

The Ray Equation. We now proceed to add the electric and magnetic

forces in each direction, working in cylindrical coordinates as before; writing E_z , H_z , and E_r , H_r , for the components in the axial and radial directions respectively, we obtain

$$F_z = e(E_z + r\dot{\phi}H_r), \quad F_r = e(E_r - r\dot{\phi}H_z), \quad F_\phi = e(\dot{r}H_z - z\dot{H}_r),$$

there being no electric force E_ϕ around the axis. Similarly we have for the respective accelerations, taking account of the centrifugal action radially,

$$a_z = \ddot{z} \quad a_r = \ddot{r} - r\dot{\phi}^2 \quad a_\phi = (1/r)d/dt(r^2\dot{\phi}) = r\ddot{\phi} + 2\dot{r}\dot{\phi}.$$

The last equation expresses the fact that the torque about the axis is equal to the rate of change of angular momentum; in the approximate treatment of the short magnetic field in section 31, only the first of the two terms involved ($r\dot{\phi}$) was used (cf. IV.5). Combining the two sets of equations we now have

$$\left. \begin{aligned} \ddot{z} &= (e/m)(\partial V/\partial z + r\dot{\phi}H_r) & (a) \\ \ddot{r} - r\dot{\phi}^2 &= (e/m)(\partial V/\partial r - r\dot{\phi}H_z) & (b) \\ d/dt(r^2\dot{\phi}) &= (e/m)(\dot{r}H_z - z\dot{H}_r). & (c) \end{aligned} \right\} \quad (IV.47)$$

Substituting in the last expression for H_z and H_r from IV.36 we get

$$\begin{aligned} d/dt(r^2\dot{\phi}) &= (e/m)[\dot{r}\partial(rA)/\partial r + rz\partial A/\partial z] \\ &= (e/m)d/dt(rA), \end{aligned}$$

the total differential of (rA) . Hence

$$r^2\dot{\phi} = (e/m)rA + C, \quad (IV.48)$$

where C is a constant of integration, depending on the initial rotation of the electron, and A is the value of the vector potential at the momentary position of the electron.† On substituting the angular velocity, as thus given, in expressions IV. 47 *a* and *b*,

$$\ddot{z} = (e/m)\left[\partial V/\partial z - \frac{1}{r}\{(e/m)rA + C\}\partial A/\partial z\right],$$

$$\text{and } \ddot{r} = (e/m)\left[\partial V/\partial r - \frac{1}{r^2}\{(e/m)rA + C\}\partial(rA)/\partial r + \frac{m}{er^3}\{(e/m)rA + C\}^2\right],$$

where H_z and H_r have again been replaced by the equivalent terms in A . These equations may conveniently be written

$$\ddot{z} = \partial/\partial z[(e/m)V - (1/2r^2)\{(e/m)rA + C\}^2] \quad (IV.49)$$

$$\text{and } \ddot{r} = \partial/\partial r[(e/m)V - (1/2r^2)\{(e/m)rA + C\}^2]. \quad (IV.50)$$

The motion of the electron in the meridional plane can therefore be represented as taking place in a potential field given by

$$\begin{aligned} Q &= (e/m)V - (1/2r^2)\{(e/m)rA + C\}^2 \\ &= (e/m)V - \frac{1}{2}r^2\dot{\phi}^2, \end{aligned} \quad (IV.51)$$

as discussed in the previous section.

† In fact, $C = r^2\dot{\phi} - \frac{e}{m}rA = r_a^2\dot{\phi}_a - \frac{e}{m}r_aA_a$, where the subscript *a* indicates values taken at the point of origin of the electron. It should be noted that this definition of C differs by the factor m/e from that used in previous books on electron optics; we prefer to follow Busch and Glaser in this respect.

Limiting the discussion to paraxial conditions, we now substitute in expression IV. 47 b for $(\partial V/\partial r)$ and H_z . The value of the former is given by taking the first two terms of the series for V_z (IV. 43), and was shown earlier to be $-(r/2)V_z''$ (II. 13); the value of the latter in terms of the axial magnetic field H_{0z} is given by the series IV. 45, of which the first term only is sufficient here. We then have

$$\ddot{r} - r\dot{\phi}^2 = -(e/2m)rV_z'' - (e/m)r\dot{\phi}H_{0z}. \quad (\text{IV. 52})$$

Applying to IV. 48 the expression for A in terms of H given earlier (IV. 44), of which the first term is again sufficient, we have

$$r^2\dot{\phi} = (e/2m)r^2H_{0z} + C.$$

For rays starting from a point on the axis outside the field, that is, with no initial angular velocity, the value of C is zero, and we can substitute at once for $\dot{\phi}$ in IV. 52:

$$\begin{aligned} \ddot{r} &= -(e/2m)rV_z'' - (e^2/2m^2)rH_{0z}^2 + (e^2/4m^2)rH_{0z}^2 \\ &= -(e/2m)r\{V_z'' + (e/2m)H_{0z}^2\}. \end{aligned} \quad (\text{IV. 53})$$

Glaser has shown that this equation holds also for the general case, when C is not zero and the object is immersed in the magnetic field (cf. section 38, p. 118); practical experience had already proved that images conforming to the normal optical rules were indeed obtained in such conditions.

Expression IV. 53 corresponds in form with those obtained earlier for the electric field (II. 14) and the magnetic field (IV. 7) acting alone. Remembering that the axial velocity $\dot{z} = \sqrt{(2e/m)V_z}$, we may write

$$\ddot{r} = \dot{z} d/dz (\dot{z} dr/dz) = 2(e/m)\sqrt{V_z} d/dz (\sqrt{V_z} dr/dz);$$

and IV. 53 then gives the differential equation for the ray path:

$$\sqrt{V_z} d/dz (\sqrt{V_z} dr/dz) = -\frac{1}{4}rV_z'' - (e/8m)rH_{0z}^2 = -Pr, \quad (\text{IV. 54})$$

where

$$P = \frac{1}{4}V_z'' + (e/8m)H_{0z}^2.$$

The radial acceleration, being negative, is therefore directed towards the axis and is proportional to the radial ordinate, so that the primary condition for image formation is fulfilled. In the general case equation IV. 54 will also contain a term relating to space charge, as well as a term in C ($= (m/2e)C^2/r^2$). When the second term in the series for H_z and A is included, as was the second term in the expression for V_z , the form of equation IV. 54 remains the same, but now

$$P = \frac{V_z''}{4} + \frac{e}{8m} \left(H_{0z}^2 - \frac{r^2}{2} H_{0z} H_{0z}' + \frac{1}{8} r^4 H_{0z}^{\prime 2} \right).$$

From IV. 48 an expression may now be obtained for the rotation of the image; we have

$$\begin{aligned} \dot{\phi} &= (e/m)rA/r^2 + C/r^2 \\ &= (e/m)H_{0z}/2 + C/r^2, \end{aligned}$$

substituting the first term of the series expression for A . The angular velocity may also be written as

$$\begin{aligned}\dot{\phi} &= (d\phi/dz)(dz/dt) = z d\phi/dz \\ &= \sqrt{\{(2e/m)V_z\}} d\phi/dz,\end{aligned}\quad (\text{IV.55})$$

if we express the axial velocity in terms of the accelerating potential. Hence

$$\sqrt{V_z} d\phi/dz = \sqrt{(e/8m)H_{0z}} + \sqrt{(m/2e)C/r^2}.$$

Setting $C = 0$, as before, we obtain for the angular rotation

$$\phi_0 = \sqrt{(e/8m)} \int_{-\infty}^{\infty} H_{0z}/\sqrt{V_z} dz. \quad (\text{IV.56})$$

For paraxial conditions, therefore, all rays starting from the object will suffer the same rotation; consequently the image will not be distorted, but will experience a rotation as a whole with respect to the object.

When the electrons are accelerated in a separate system before the magnetic lens, and there is no electrostatic field within the latter, the above expressions IV.54 and IV.56 reduce to the form of those derived for the simple magnetic lens (IV.8 and IV.13).

Use of Special Solutions. The general differential equation (IV.54) for the path of a ray in combined electric and magnetic fields will have as many solutions as there are possible paths from object to image. If two special solutions are known, then any other path may be expressed in terms of them. As shown in section 26 (p. 65), two particular solutions r_1 and r_2 are related at every value of z by the expression

$$r_2 r'_1 - r_1 r'_2 = C_{12}/\sqrt{V_z}, \quad (\text{IV.57})$$

where C_{12} is a constant of integration depending on the two paths chosen, and V_z is the potential at the axial point at which particular values of the radial distances (r_1, r_2) and inclinations (r'_1, r'_2) of the rays are measured. Any other solution r_x of the ray equation can be similarly combined with r_1 and r_2 , in turn, so that

$$\left. \begin{aligned} r_x r'_1 - r_1 r'_x &= C_{1x}/\sqrt{V_z} \\ r_2 r'_x - r_x r'_2 &= C_{x2}/\sqrt{V_z} \end{aligned} \right\}. \quad (\text{IV.58})$$

Multiplying these two equations by r_2 and r_1 , respectively, and adding,

$$r_x(r_2 r'_1 - r_1 r'_x) = (r_2 C_{1x} + r_1 C_{x2})/\sqrt{V_z}.$$

On division by IV.57 this becomes

$$\left. \begin{aligned} r_x &= r_1 C_{x2}/C_{12} + r_2 C_{1x}/C_{12} \\ &= r_1 c_1 + r_2 c_2 \end{aligned} \right\}, \quad (\text{IV.59})$$

where c_1 and c_2 are again constants characteristic of the particular pair of solutions assumed known.

In section 26 two solutions were chosen corresponding to paths parallel

to the axis in the object and image space respectively, in order to locate the focal points and obtain a relation (III.28) between them. For the later discussion of aberrations, however, it is convenient now to take a pair of rays (r_α, r_β) specified by the following values in the object plane:

$$\left. \begin{aligned} r_{\alpha(0)} &= 0; & r'_{\alpha(0)} &= 1 \\ r_{\beta(0)} &= 1; & r'_{\beta(0)} &= 0 \end{aligned} \right\}, \quad (\text{IV.60})$$

the ray r_β being parallel to the axis at unit distance and r_α intersecting the axis in the object plane at unit inclination, so that its radial distance in the aperture ($r_{\alpha B}$) will equal the object distance (p). Writing for the unknown solution r_x the value r_0 in the object plane, we have from equations IV.57 and IV.58,

$$\begin{aligned} C_{12} &= \sqrt{V_0}; & r_0 &= C_{1x}/\sqrt{V_x}, \text{ or } C_{1x} = r_0\sqrt{V_0}; \\ r'_0 &= C_{x2}/\sqrt{V_0}, \text{ or } C_{x2} = r'_0\sqrt{V_0}; \end{aligned}$$

where V_0 is the axial potential in the object plane. Substituting in IV.59 we have

$$r_x = r_\alpha r'_0 + r_\beta r_0, \quad (\text{IV.61})$$

giving an equation for the value of the unknown ray at every axial position, in terms of the values of the special solutions at that point and the parameters of r_x in the object plane. Similarly it may be shown that the local value of the inclination of the ray is given by

$$r'_x = r'_\alpha r'_0 + r'_\beta r_0. \quad (\text{IV.62})$$

These two equations completely describe the path of any ray in terms of the two special solutions r_α and r_β .

The Focal-length Equation. An expression for the focal length of the combined lens may be obtained in the same manner as for the electrostatic field alone (section 26). By integration of IV.54, writing P as there defined,

$$\sqrt{V_n} r'_n - \sqrt{V_0} r'_0 = - \int_0^n Pr/\sqrt{V_x} dz,$$

where the subscript 0 refers to the origin and the subscript n to the image space. Allowing the entrant beam to be parallel to the axis, $r'_0 = 0$, and

$$r'_n = -(1/\sqrt{V_n}) \int_0^n Pr/\sqrt{V_x} dz.$$

Then the (second) focal length f_2 is given by

$$1/f_2 = -r'_n/r_0 = \{1/(r_0\sqrt{V_n})\} \int_{-\infty}^{\infty} Pr/\sqrt{V_x} dz, \quad (\text{IV.63})$$

taking the limits from infinity to infinity, since no field exists outside the lens. As r is contained within the integrand this expression is not generally susceptible of treatment and an expression for f can only be obtained in the form of a series of terms, of the same nature as given in III.23 for the

electrostatic lens, but with the quantity (P) instead of ($V''/4$) under the integral sign:

$$\begin{aligned} \frac{1}{f} &= \frac{1}{\sqrt{V_n}} \int_{-\infty}^{\infty} \frac{P}{\sqrt{V_z}} dz - \frac{1}{\sqrt{V_n}} \int_{-\infty}^{\infty} \frac{P}{\sqrt{V_z}} \int_{-\infty}^z \frac{1}{\sqrt{V_n}} \int_{-\infty}^z \frac{P}{\sqrt{V_z}} dz dz dz \\ &+ \frac{1}{\sqrt{V_n}} \int_{-\infty}^{\infty} \frac{P}{\sqrt{V_z}} \int_{-\infty}^z \frac{1}{\sqrt{V_n}} \int_{-\infty}^z \frac{P}{\sqrt{V_z}} \int_{-\infty}^z \frac{1}{\sqrt{V_n}} \int_{-\infty}^z \frac{P}{\sqrt{V_z}} dz dz dz dz dz - \dots \quad (\text{IV.64}) \end{aligned}$$

Making the same simplifying assumption as previously, that the radial distance of the electron within the lens is constant and equal to r_0 , we obtain from IV.63 an approximate relation for f :

$$\begin{aligned} 1/f &= (1/\sqrt{V_n}) \int_{-\infty}^{\infty} P/\sqrt{V_z} dz \\ &= \{1/(4\sqrt{V_n})\} \int_{-\infty}^{\infty} V_z''/\sqrt{V_z} dz + \{e/(8m\sqrt{V_n})\} \int_{-\infty}^{\infty} H^2/\sqrt{V_z} dz \\ &= 1/f_e + 1/f_m, \end{aligned}$$

as will at once be evident on referring to the expressions obtained for the focal length of the simple electrostatic lens f_e (III.23) and of the isolated magnetic lens f_m (IV.10). It thus appears that, to a first approximation, the superimposed fields may be considered as each giving rise to the lens which it would form if present alone, the total power being given by the simple sum of the two components treated as superimposed thin lenses. This result is of great value when using lenses employing both electrostatic and magnetic fields; it might, of course, have been predicted from the fact that the electric and magnetic forces on the electron are simply additive.

An alternative treatment is also due to Scherzer, and is followed by Picht. Writing IV.54 in a form corresponding to that given earlier (II.16) for the simple electrostatic lens,

$$Vr'' + \frac{1}{2}V'r' = -\frac{1}{4}rV'' - (er/8m)H^2,$$

he introduces the substitution

$$R = rV^{\frac{1}{2}}. \quad (\text{IV.65})$$

As a result,

$$R'' = -R\left[\frac{3}{16}(V'/V^2) + eH^2/8mV\right] = -RT. \quad (\text{IV.66})$$

Corresponding to IV.63 we then obtain, in the light of IV.65,

$$1/f_2 = -r'_n/r_0 = \left(\frac{V_0}{V_n}\right)^{\frac{1}{2}} R'_n/R_0 = \left(\frac{V_0}{V_n}\right)^{\frac{1}{2}} \int_{z_0}^{z_1} T dz, \quad (\text{IV.67})$$

on substituting for R from IV.66. This integral again cannot be evaluated directly, but must be solved by successive approximation on the lines of

expression IV.64. It has the advantage, however, of containing only the first and not the second derivative of the potential, a matter of great practical importance since the former may be determined with much greater accuracy than the latter from an experimental investigation of the potential distribution in a lens. For theoretical purposes, and especially in the discussion of aberrations of the image, it is preferable to employ the method given first above.

FURTHER READING

1. Solenoidal Focusing:

- Tricker, *Proc. Camb. Phil. Soc.* **22**, 454, 1924.
 Farnsworth, *J. Franklin Inst.*, **218**, 411, 1934.
 Witcher, *Phys. Rev.* **60**, 32, 1941.

2. Short Lens:

- Busch, *Ann. Phys.* **81**, 974, 1926.
 Bouwers, *Physica*, **4**, 200, 1937 (in English).
 Preston, *J. Sci. Inst.* **21**, 205, 1944.

3. Form of Field:

- Klemperer, *Phil. Mag.* **20**, 545, 1935.
 Klemperer and Miller, *J. Sci. Inst.* **26**, 121, 1939.
 Dosse, *Z. Phys.* **117**, 437, 1941.
 Siegbahn, *Ark. Mat. Astr. Fys.* **30**, 1, 1943 (in English).
 Wallauschek and Bergmann, *Z. Phys.* **94**, 329, 1935.

4. Immersion Objective:

- Ruska, *Z. Phys.* **89**, 90, 1934.
 Glaser, *Z. Phys.* **117**, 285, 1941.
 Marton and Hutter, *P.I.R.E.* **32**, 3, and 546, 1944.

5. Ray-tracing:

- Siday, *Proc. Phys. Soc.* **54**, 266, 1942.
 Morton and Ramberg, *Physics*, **7**, 451, 1936.
 Goddard and Klemperer, *Proc. Phys. Soc.* **56**, 378, 1944.

6. Magnetic Vector Potential:

- Pender and Warren, *Electric Circuits and Fields* (Appendix), 1943.

7. Combined Electrostatic and Magnetic Fields:

- Glaser and Lammle, *Ann. Phys.* **40**, 367, 1941.

CHAPTER V

ABERRATIONS OF THE IMAGE

37. *Introduction*

THE discussion of image formation has been confined so far to paraxial conditions, in which contributory rays make with the axis an angle so small that its value (in radian measure) is not significantly different from that of the sine; for angles less than 10° the difference is negligible for normal purposes. In such a case all rays from a given point of the object come together again in a single point of the image; the object gives a true image, which is then said to be aberration-free. The cone of rays gathered by the lens must thus be limited to a semi-aperture of less than about $\pi/20$, and the brightness of the image relative to the object is correspondingly restricted. For many purposes it is necessary to gather a larger proportion of the emitted electrons, and a bigger aperture is used; those paths which make slightly greater angles with the axis than the paraxial rays are termed zonal, and those farther out still are termed marginal. For these rays the value of the sine of the angle approximates closely to that obtained from the first two terms in the series

$$\sin \theta = \theta - \theta^3/3! + \theta^5/5! - \dots;$$

the error is less than 1 per cent. up to angles of one radian. When, as in optics, still larger apertures are employed, the third term of the expansion must also be taken into account.

Under these conditions the simple expressions evolved earlier for the relation between object and image no longer apply. It is found that the image is more or less seriously distorted with respect to the object: electrons from the same point on the object intersect the image plane in different points, and the image plane itself may be curved. The investigation of these differing paths may be conducted analytically by employing the first two terms of the above sine expansion, and using trigonometrical or Hamiltonian (wave) methods. The resulting theory is known as the third-order theory, and yields a comprehensive equation for the divergence of an image point from that predicted by the first-order theory. This proves to be too complicated for practical discussion; and is conveniently resolved into a set of terms each of which characterizes a particular form of aberration of the image. In special conditions each aberration can be discussed mathematically and observed in practice separate from the others. In general the image will suffer from varying proportions of all the aberrations, and consequently will exhibit greater or less confusion.

The form of the relationship between the radial coordinates of the electron path and the different aberration terms can be deduced by consideration of the symmetry conditions. Suppose an object plane; with

axes x_0, y_0 , to intersect the axis (z - z) of the optical system orthogonally in a point O (Fig. 63). An image will be predicted by the Gaussian first-order theory as being formed in a second plane (x_i, y_i) at right angles to the axis at a distance z_i from O . For the purpose of the present discussion we may represent the focusing system by a lens-plane, x_L, y_L , parallel to the object plane. Any point P in the object plane, at a radial distance r_0 from the axis, will be imaged in a point P_i in the image plane, at a distance r_i from the axis, according to the paraxial theory. A possible path for such a ray is indicated by the full line in Fig. 63; it will intersect the

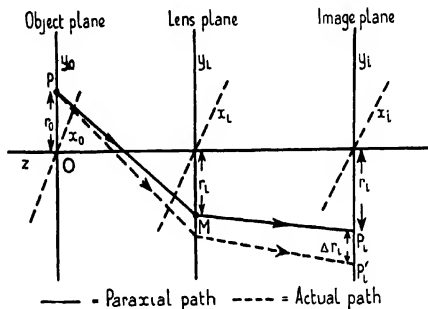


FIG. 63. Idealized ray-paths in lens, showing radial aberration (Δr_i).

lens-plane in some point M at a distance r_L from the axis. In practice there will exist a maximum possible value for r_L , determined by the physical apertures comprised in the lens. Now the actual path followed by a ray from P will differ from that shown, PMP_i , by an amount depending on how much greater an angle PM makes with the axis than is allowed by paraxial conditions. The broken line indicates the typical deviation that results, leading to an image point P'_i at a distance r'_i from the axis. The displacement of P'_i from P_i will be the total lateral aberration,

$$\Delta r_i = r'_i - r_i.$$

The value of Δr_i may be expressed in the form of a power series of r_0 and r_L containing, in the third-order theory, terms up to and including those of the third power, thus:

$$\Delta r_i \rightarrow r_0; r_L; r_0^2; r_L^2; r_0 r_L; r_0^3; r_L^3; r_0^2 r_L; r_0 r_L^2; r_L^3.$$

But when r_0 and r_L are small, $r_i = r'_i$, and hence the difference, Δr_i , must be independent of the first power of r_0 and r_L . Also on account of the cylindrical symmetry of the usual lens systems, the aberration may be shown to be independent of even-order terms as well, and there remain only the third-order terms. We may thus write

$$\Delta r_i \rightarrow r_0^3; r_0^2 r_L; r_0 r_L^2; r_L^3.$$

Each term may then be related to a particular form of divergence of the image point from the paraxial point, P_i ; in light optics these are classified as five errors of the image:

Distortion (first term)	
Curvature of the field	} (second term)
Astigmatism	
Coma (third term)	
Spherical aberration (fourth term).	

In general there will be rotational as well as radial aberration of the image; that is, the point P'_i will lie not on the y_i -axis but on a radius rotated through some angle χ from it. On this account the second term in the above relation proves to consist of two parts, so that light optical theory recognizes five aberration terms. By more detailed discussion along these lines Ramberg was able to obtain the exact interrelation of the coefficients of the several terms. It will be noted that for an object point on the axis ($r_0 = 0$) all terms except the last vanish, so that the image is then afflicted only by spherical aberration.

In electron optics the same number of aberrations are found to exist in the electrostatic field acting alone, as is to be expected from its isotropic nature. For combined electric and magnetic fields Glaser has shown that a series of nine terms results from the Hamiltonian treatment, specifying eight independent third-order errors. Three new forms of aberration result from the rotation of the electrons in the magnetic field, and to these are given the names:

Anisotropic astigmatism,
 Anisotropic coma,
 Anisotropic distortion,

since they arise from the essential anisotropy of the field. They produce rotational effects in the normal aberrations of corresponding type (cf. Fig. 71). This method of approach, which makes use of the Hamiltonian 'eiconal', is indicated in the Appendix (p. 249). The alternative trigonometrical method was developed by Scherzer, using the conceptions employed in earlier chapters for the discussion of electron trajectories. There results in this case a set of six terms descriptive of the eight errors of electromagnetic lenses, some of them containing two of Glaser's coefficients. An outline of Scherzer's treatment will be given here, and reference should be made to the original papers for further details. Each aberration will then be separately described and its importance in electron optics and the possibilities of correction briefly discussed; in each case the aberration integral will be quoted, without derivation. The discussion will assume the incident electron beam to be homogeneous as regards velocity. The variation in the position of the image that is found for electrons of varying velocities (chromatic aberration) requires separate treatment (section 43). At the outset it must be said that the spherical and chromatic errors constitute the major aberrations of electron lenses and that so far very little experimental investigation of other errors has been undertaken, in comparison with the development of their theoretical treatment.

Before commencing the mathematical discussion it will be as well to

indicate its limitations and advantages, as compared with the light optical problem. Owing to the difficulty of making other than spherical surfaces in glass optics, the classical theory of aberrations has been concerned to determine the deviation of an actual ray from the path given by paraxial considerations for its passage through one or more such surfaces. The number of refracting boundaries is usually small and their form comparatively simple; the lack of a continuously varying refractive index, however, makes the problem difficult of exact mathematical analysis. In electron optics, on the other hand, this latter condition is fulfilled by the very nature of the field and refraction occurs at all points, so that the trajectory is a smooth curve; but the form of the field, unfortunately, is not at all simple and can rarely be expressed by a tractable mathematical function. Theoretical discussion has naturally tended to follow the light optical method of finding expressions for the shift of the image point from the position given by paraxial theory. In effect the starting-point is the same as in optics: the focusing action of the lens field is taken as proportional to the distance of the electron from the axis in passing through the lens. The ray equation is then brought into closer correspondence with the real field by introducing secondary terms, and the aberration found as the difference between the path as given by this equation and the paraxial form of it. As in light optics the final expressions are cumbersome, and it will not be surprising if progress in the immediate future is made rather as a result of detailed ray-tracing than by the application of analysis.

It is, however, also possible that a more fruitful theoretical approach may be found in a new direction: after developing and possibly simplifying the equations for individual aberrations, to determine the form of the electrodes which will give a field for which one or more errors have a minimum value. In light optics this method is not practised owing to its technical limitation to surfaces of a particular shape, and, on the other hand, to the availability of a variety of media of widely differing properties from which a corrected combination may be devised. In electron optics, as pointed out in the Introduction, the field has essentially the same properties in all lenses whether electrostatic or magnetic, and complete correction in the optical sense of the term is impossible. On the other hand, fields of almost any form can be established by varying the shape and separation of the electrodes or pole-pieces and by the controlled use of space charge. Mention will be made of a number of attempts to deduce the form of fields of minimum aberration, from which the boundary conditions (i.e. the form of the electrodes) may be found by calculating the distribution of equipotentials, any set of which can then be replaced by suitably charged conducting surfaces of the same shape, without changing the field distribution. Positive progress has so far been small, but it is obviously of value to explore the differences as well as the similarities between light and electron optical systems; it may, indeed, prove more fruitful to deal with the total aberration equation rather than to concentrate on individual errors in turn.

The possibility of employing distributed space charge to correct lenses, or to produce a perfect lens of new type, is discussed in section 44; in glass optical terms it is equivalent to finding a medium with refractive index decreasing with distance from the axis, or to introducing varying concentrations of one medium into the body of another. A simple expression for the form of the perfect lens is obtainable, but its translation into practice is obstructed by the difficulty of controlling the distribution of space charge.

*38. General Treatment of Aberration

Similar aberrations are found in both electrostatic and magnetic lenses, although the relative amounts of the various distortions may differ considerably. We shall therefore take the general case of combined electric and magnetic fields and, proceeding on lines indicated at the end of the previous chapter (section 36), derive an expression for the total aberration. Further development leads to a number of terms each characteristic of a particular form of divergence from the true image calculated by the first-order theory.

In obtaining expression IV.54 it was assumed that the constant C (from IV.48) had the value zero, which is equivalent to limiting discussion to rays starting from a point on or near the axis and having no initial component of angular velocity about it. In order to consider the contribution of all rays forming the image of an off-axis point, account must now be taken of skew rays, for which C is finite. The practical result is that the image point formed by skew rays may be rotated about the axis by an angle that differs from the value ϕ_0 given by expression IV.56, for which it was assumed that $C = 0$. Let the difference in rotation be the angle χ , so that the true value of the rotation of an electron after travelling a distance z is given by

$$\phi = \sqrt{\left(\frac{e}{8m}\right)} \int_0^z \frac{H_z}{\sqrt{V}} dz + \chi. \quad (\text{V.1})$$

We shall now show that χ is zero for an image formed by paraxial rays.

It proves to be convenient to combine in complex notation the radial and rotational parts of the trajectory and to obtain a general solution. For this purpose we define a parameter u such that

$$u = x + iy = r \cos \chi + ir \sin \chi = re^{i\chi}, \quad (\text{V.2})$$

when the electron is at a distance r from the axis and has suffered a rotation χ from the position predicted by the paraxial equation, that is, from the rotating meridian plane which contains the x -axis and to which the y -axis is perpendicular. Then when $C \neq 0$, we must write in place of IV.54 and IV.56

$$\sqrt{V} \frac{d}{dz} \left(\sqrt{V} \frac{dr}{dz} \right) = -Pr + \frac{m}{2e} \frac{C^2}{r^3}, \quad (\text{V.3})$$

$$\sqrt{V} \frac{d\phi}{dz} = \sqrt{\left(\frac{e}{8m}\right)} H + \sqrt{\left(\frac{m}{2e}\right)} \frac{C}{r^2}, \quad (\text{V.4})$$

where P has, as before, the value

$$P = \frac{1}{4}V'' + (e/8m)H^2,$$

it being understood that V and H are the axial values. From V.1 and V.4 we have at once

$$d\chi/dz = \sqrt{\left(\frac{m}{2eV}\right) \frac{C}{r^2}} = \sqrt{\left(\frac{m}{2eV}\right) \frac{C}{u^2}} e^{2i\chi}. \quad (\text{V.5})$$

Putting $r = ue^{-i\chi}$ in V.3 we get a differential equation for u :

$$\sqrt{V} \frac{d}{dz} \left(\sqrt{V} \frac{du}{dz} \right) = -Pu, \quad (\text{V.6})$$

which will have a number of solutions corresponding to the possible ray paths from object to image. A particular path may be represented (cf. IV. 59) in terms of two linearly independent solutions, r_α and r_β , and two arbitrary constants c_α and c_β ,

$$u(z) = c_\alpha r_\alpha + c_\beta r_\beta, \quad (\text{V.7})$$

such that V.6 is satisfied by both $u = r_\alpha$ and $u = r_\beta$. Any two representative rays may be chosen for r_α and r_β .

In the plane $z = z_a$, which may conveniently be the object plane, we set $r_\alpha = 0$ and $r'_\alpha = 1$, with the result that the first term in V.7 vanishes. If we then set $r_\beta = 1$ and $r'_\beta = 0$ in the same plane, we have from V.1 and V.7,

$$c_\beta = u_a = r_a; \quad c_\alpha = u'_a = r'_a + ir_a \chi'_a = r'_a + ir_a \left\{ \phi'_a - \sqrt{\left(\frac{e}{8mV_a}\right)} H_a \right\},$$

where r_a is the radial distance of the origin of the electron in the plane $z = z_a$, where $\chi = 0$. Now if the fields are sufficiently extensive, there will be another plane ($z = z_b$) in which again $r_\alpha = 0$, and therefore

$$u_b = c_\beta r_{\beta(b)} = r_a r_{\beta(b)}.$$

This solution is real, and consequently in this plane we must also have $\chi = 0$ and $r_b = u_b = r_a r_{\beta(b)}$; so that both r and ϕ prove to be independent of the initial tangential direction of the rays from the object, and the conditions for true image formation by a paraxial beam are fulfilled. The magnification is given by $r_b/r_a = r_{\beta(b)}$. Then the general form of the solution satisfying the initial conditions is

$$u = c_\alpha r_\alpha + c_\beta r_\beta = r_\alpha \left[r'_a + ir_a \left\{ \phi'_a - \sqrt{\left(\frac{e}{8mV_a}\right)} H_a \right\} \right] + r_a r_\beta, \quad (\text{V.8})$$

and we have, from V.1 and V.2,

$$r_p = |u| = r_a \sqrt{\left[\left(r_\beta + \frac{r'_a}{r_a} r_\alpha \right)^2 + \left\{ \phi'_a - \sqrt{\left(\frac{e}{8mV}\right)} H_a \right\}^2 r_\alpha^2 \right]} \quad (\text{V.9})$$

$$\text{and.} \quad \phi_p = \sqrt{\left(\frac{e}{8m}\right)} \int_0^s \frac{H}{\sqrt{V}} dz + \arg u$$

$$= \sqrt{\left(\frac{e}{8m}\right)} \int_0^s \frac{H}{\sqrt{V}} dz + \tan^{-1} \frac{\phi_a - \sqrt{(e/8mV)} H_a}{r_\beta/r_\alpha + r'_a/r_a}. \quad (\text{V.10})$$

The suffix p has been added to both r and ϕ in these expressions in order to indicate that the solutions apply for a paraxial path only.

It is now necessary to extend the treatment to object points so far off the axis that the potential and field strength are no longer given with sufficient accuracy by the first terms of the series obtained from the axial values (IV.43 and IV.45). By taking the first three terms for V and the first two for H we get for the equivalent potential in the meridian plane (cf. IV.51):

$$\frac{m}{e}Q = -\frac{e}{2m}\left[\frac{1}{2}Hr - \frac{1}{16}H''r^3 + \frac{mC}{er}\right]^2 + V - \frac{1}{4}V''r^2 + \frac{1}{84}V^{(4)}r^4.$$

Neglecting squares of derivatives of H and V , and higher terms of r , we have for the axial and rotational components of the trajectory:

$$\sqrt{V_z} \frac{d}{dz} \left(\sqrt{V_z} \frac{dr}{dz} \right) = -Pr + \frac{m}{2e} \frac{C^2}{r^3} + \frac{1}{32} V^{(4)} r^3 + \frac{e}{16m} H' \left(H + \frac{mC}{r^2} \right) r^3, \quad (\text{V.11})$$

$$\text{and} \quad \sqrt{V_z} \frac{d\phi}{dz} = \sqrt{\left(\frac{e}{8m} \right) H - \frac{1}{16} \sqrt{\left(\frac{e}{12m} \right) H'' r^2} + \sqrt{\left(\frac{m}{2e} \right) \frac{C}{r^2}}, \quad (\text{V.12})$$

where the value of V_z is given by the energy relation

$$\begin{aligned} V_z &= (m/2e)v_z^2 = V - \frac{1}{4}V''r^2 - Vr'^2 - Vr^2\phi'^2 \\ &= V - \frac{1}{4}V''r^2 - Vr'^2 - \frac{e}{2m} \left(\frac{1}{2}Hr + \frac{mC}{e} \frac{1}{r} \right)^2. \end{aligned} \quad (\text{V.13})$$

These expressions will in principle now give the trajectory for any non-paraxial ray. Following optical practice we choose to evaluate the difference between the radial distance of such a ray and that of a corresponding paraxial ray at the same z -plane, and the similar difference in rotation about the axis. We write

$$r = r_p + \rho; \quad \phi = \phi_p + \delta,$$

with r_p and ϕ_p as given by V.9 and V.10 for the paraxial ray. The small quantities ρ and δ will then specify the degree to which the given ray departs from the path of the paraxial ray at any value of z . Expressions for the aberration factors are then obtained from V.11 and V.12:

$$\sqrt{V} \frac{d}{dz} \left(\sqrt{V} \frac{d\rho}{dz} \right) = - \left(P + \frac{3}{2} \frac{m}{e} \frac{C^2}{r_p^3} \right) \rho + B, \quad (\text{V.14})$$

$$\sqrt{V} \frac{d\delta}{dz} = \frac{2e}{m} \left(D - \frac{mC}{er_p^2} \right) \rho, \quad (\text{V.15})$$

where B , C , and D are functions of V and H and their derivatives with respect to z .

Equations V.14 and V.15 are linear in ρ and δ , and can be solved once we know the paraxial trajectory, given by r_p and ϕ_p . Since the latter already satisfy the initial conditions that have also to be satisfied by the non-paraxial ray, r and ϕ , then ρ and δ must vanish at the plane $z = z_0$, as is

implicit in the formulation of the problem. With these initial conditions for ρ and δ , and writing $r^2 = u\bar{u}$, where $u = re^{ix}$ as before, and $\bar{u} = re^{-ix}$, a solution of V.14 and V.15 may be obtained by lengthy reasoning, which it is unnecessary to reproduce here, in the complex form:

$$\rho + i\delta = \frac{r_\alpha}{\sqrt{V_\alpha}} e^{-ix} \int_{z_\alpha}^z \frac{B_1 r_\beta}{\sqrt{V}} e^{ix} dz - \frac{r_\beta}{\sqrt{V_\alpha}} e^{-ix} \int_{z_\alpha}^z \frac{B_1 r_\alpha}{\sqrt{V}} e^{ix} dz, \quad (\text{V.16})$$

r_α and r_β being the paraxial solutions of V.7. In the image plane, where $r_\alpha = 0$ and $r_\beta e^{-ix} = M$, the magnification, this becomes

$$\rho + i\delta = -\frac{M}{\sqrt{V_\alpha}} \int_{z_\alpha}^z \frac{B_1 r_\alpha}{\sqrt{V}} e^{ix} dz; \quad (\text{V.17})$$

giving an expression for the total aberration with respect to the paraxial path. In these expressions B_1 is given by

$$B_1 e^{ix} = u' \left[V\omega_1 + i \sqrt{\left(\frac{eV}{2m}\right)} \left(\omega_1 H - \frac{r^2}{4} H'' \right) \right] + u \left[\frac{r^2}{32} V^{17} - \frac{\omega_1}{2} V'' + \frac{er^2}{16m} HH'' - \frac{e\omega_1}{4m} H^2 + i \sqrt{\left(\frac{eV}{8m}\right)} \left(\omega_1 H' + \omega_1' H - \frac{r^2}{8} H^{17} \right) \right]; \quad (\text{V.18})$$

where

$$\omega_1 = \frac{r^2 V''}{8V} + \frac{r'^2}{2} + \frac{r^2}{2} \left(x' + \sqrt{\left(\frac{e}{8mV}\right)} H \right)^2$$

$$= \frac{r^2 V''}{8V} + \frac{1}{2} u' \bar{u}' + \frac{er^2 H^2}{16mV} + \frac{i}{8} \sqrt{\left(\frac{2e}{mv}\right)} H(u\bar{u}' - \bar{u}u'). \quad (\text{V.19})$$

Although it contains all the different aberrations, it is more convenient to resolve V.17 into a number of terms, each characteristic of one particular error. The general nature only of this procedure will be indicated here, and the final result quoted.

Attention is now directed to rays whose path is conditioned by some suitable aperture within the lens system; this may be an actual physical diaphragm, or it may be the virtual 'exit-pupil' of optical theory. The two fundamental solutions, combined in the quantity u , will now be r_α and r_γ , where r_α is defined as before and r_γ has the value 1 in the object plane and the value zero in the aperture. The subscript B will be used to denote quantities measured in the aperture plane, $z = z_B$. Then

$$r_\gamma = r_\beta - \frac{r_{\beta B}}{r_{\alpha B}} r_\alpha,$$

and, substituting V.4 in the expression for u given earlier (V.8), we have

$$u = r_\alpha r_\gamma + u_B \frac{r_\alpha}{r_{\alpha B}}, \quad u_B = r_\alpha r_{\beta B} + r'_\alpha r_{\alpha B} + i C \frac{r_{\alpha B}}{r_\alpha} \sqrt{\frac{m}{2eV_\alpha}},$$

$$\bar{u} = r_\alpha r_\gamma + \bar{u}_B \frac{r_\alpha}{r_{\alpha B}}.$$

Remembering that $r^2 = u\bar{u}$, we may now expand $B_1 e^{ix}$ and ω_1 , and transform the expression for the total aberration (V.17) into a series of terms for the third-order errors; the result is

$$\rho + i r \delta = \alpha r_a^2 + \beta r_a^2 u_B + \gamma r_a^2 \bar{u}_B + \delta r_a u_B^2 + \epsilon r_a u_B \bar{u}_B + \zeta u_B^2 \bar{u}_B. \quad (\text{V.20})$$

The coefficients α, β, \dots , etc., are in the form of integrals (with respect to z) of some complexity and will be given under the discussion of the separate aberrations.

All the quantities in u in V.20 refer to the paraxial path; if the form of the field is known, they and the coefficients may be calculated and the value of each aberration determined for any non-paraxial path whose origin (r_a) is given. For $r_a = 0$ it will be seen that all terms vanish except the last (ζ) which relates to the spherical aberration; this, therefore, is the only error to which rays from a point on the axis are subject, in third-order theory. The first term (α) determines the distortion, including the spiral distortion which is characteristic of the magnetic field. The second term (β) relates to the curvature of the image plane, the third (γ) to the astigmatism, the fourth and fifth (δ and ϵ) to coma, including the anisotropic coma due to the magnetic field. Diels and Wendt have shown experimentally that each of these errors do indeed occur in electron lenses. Detailed experimental investigation, however, has so far been almost entirely confined to spherical aberration, which is the most pronounced error in both electric and magnetic fields. Less progress has been made in either experimental or theoretical treatment of the conditions in which correction of other individual aberrations might be achieved.

Voit has investigated analytically the conditions under which it may be possible to reduce them to a minimum or to eliminate them entirely. He finds that all, except spherical aberration and anisotropic distortion, can be separately caused to vanish by the use of special apertures and/or potential fields of a particular form. The use of auxiliary lenses allows of all errors, without exception, being reduced collectively to any desired limit for a given magnification and aperture. The conclusions apply to strong as well as weak lenses, but considerable technical difficulties stand in the way of their practical realization.

The above discussion has assumed the incident beam of electrons to be homogeneous, and therefore no term appears for the chromatic aberration, which arises from non-uniformity of velocity. From the form of the expressions obtained it is clear that electrons moving at different speeds will be refracted differently in the field, and each of the third-order aberrations will be affected accordingly. It is convenient, therefore, to treat the chromatic error as a separate problem (see section 43).

*39. Spherical Aberration

An optical system is said to suffer from spherical aberration when rays incident at varying radial distances are focused to different points on the axis. When the incident beam is parallel to the axis, Fig. 64 shows the

most frequent form of the aberration, in which rays incident in the outer parts of the lens (marginal rays) are brought to a focus at a point (F_m) closer to the lens than the focus (F_p) for a paraxial pencil; intermediate (zonal) rays will form a focus (F_z) within these two extremes. The aberration is said to be positive when the marginal rays intersect the axis nearer to the lens than the true or paraxial focus, as depicted; this condition (ζ having a negative sign) arises with lenses of positive power, whereas negative lenses form the marginal focus farther away than F_p , ζ having a

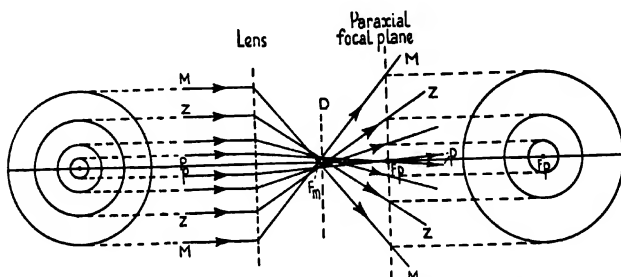


FIG. 64. Spherical aberration of an initially parallel beam: F_p = paraxial focus; F_m = marginal focus; D = disk of least confusion.

positive value. The form of the last term in the equation for the total aberration (V.20),

$$\rho_s = \zeta u_B^2 \bar{u}_B, \quad (\text{V.21})$$

shows that the radial effect (ρ_s) of the spherical aberration in the image plane is proportional to the third power of the aperture, since both u_B and \bar{u}_B depend on the radial distance of the electron in the limiting aperture of the system. A series of concentric and parallel incident pencils, seen in cross-section to the left of the diagram, will form in the (Gaussian) focal plane a series of rings whose diameter increases rapidly with the aperture (shown to the right). Alternatively, a point object O on the axis will give rise to a similar set of rings in a plane perpendicular to the axis through the paraxial image point (I , Fig. 65), but the rings will now form a continuous set, observable as a patch of light the intensity of which rapidly diminishes with distance from the axis. In both cases there will exist, at some point between F_p (or I) and the lens, a minimum cross-sectional area of the beam, the position of which may be determined by moving a screen along the axis. It is termed the disk of least confusion, being the smallest spot formed by the imaging system, and is indicated by D in the diagrams.

Experimental Investigation. In theoretical discussion the degree of spherical aberration is described, as above, in terms of the radial displacement (ρ_s) of the actual image point from the Gaussian image point, in the plane determined by the latter. For practical purposes it is easier to measure either the diameter (d) of the disk of least confusion or the separation of the marginal and paraxial focal points for an initially parallel

beam. The latter method is usually preferred, the distance $F_p F_m$ being then termed the longitudinal spherical aberration. Klemperer and Wright made measurements of this quantity for the electrostatic lenses investigated in the experimental system described in Chapter III (p. 50). The pepper-pot diaphragm in the path of the incident rays was adjusted to admit pencils of various radii in turn, and the position of the focus determined each time by direct observation on the movable screen. By this

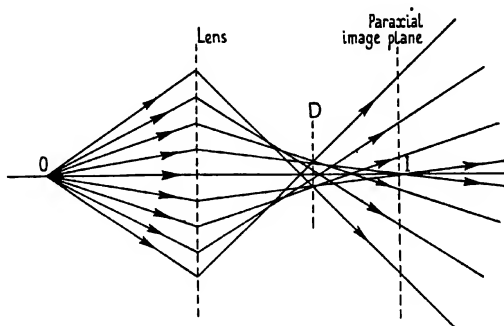


FIG. 65. Spherical aberration of an initially divergent beam:
 D = disk of least confusion.

means the set of curves shown in Fig. 66 was obtained, showing the variation in focal length with voltage ratio for different apertures of the incident beam, expressed as fractions (y/R) of the radius R of the lens electrodes. For a given voltage ratio aberration curves of the type used in light optics were then drawn, in which the focal length is shown in its dependence on the aperture of the beam (Fig. 67). It will be seen that for both accelerating and decelerating voltage ratios the focal length decreases with the initial distance of the beam from the axis: the spherical aberration is positive, as is usual with converging lenses. The degree of aberration, however, is less when the electrons are being accelerated than in the reverse case. This difference is in part due to the fact that the electrons when being accelerated experience a radial velocity that is always towards the axis; whereas when passing through the lens in the opposite direction the radial velocity is away from the axis for more than half the trajectory, carrying them into regions where the aberration factor has larger values, since it depends on the third power of the aperture. The spherical aberration is also less for high than for low accelerating ratios; the difference in practical effect is small, however, since it is the ratio of aberration to focal length which determines the clarity of the image.

Ray-tracing has shown that the outer parts of the electrostatic field contribute most to the aberration; the low-voltage (converging) side produces positive aberration, and the high-voltage (diverging) semi-lens negative aberration. It is not yet clear whether with the present form of lens construction it may be possible to balance the aberration contributions

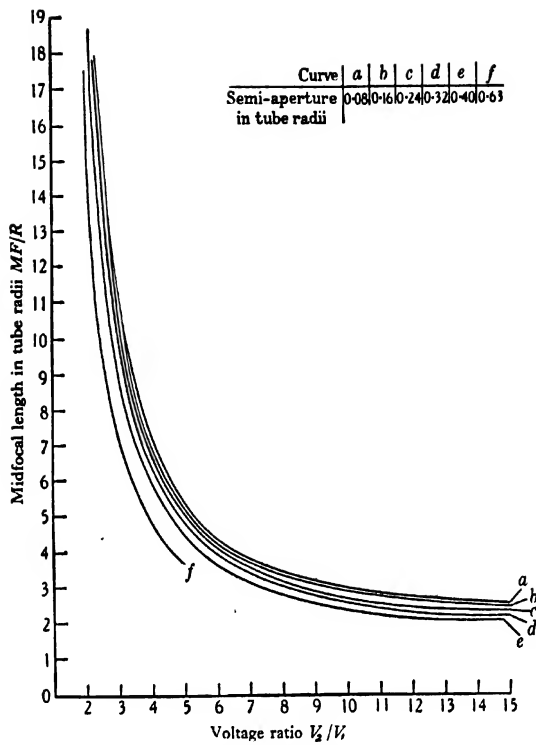


FIG. 66. Spherical aberration of two-cylinder lens: experimental values for mid-focal length at various apertures. (Klemperer.)

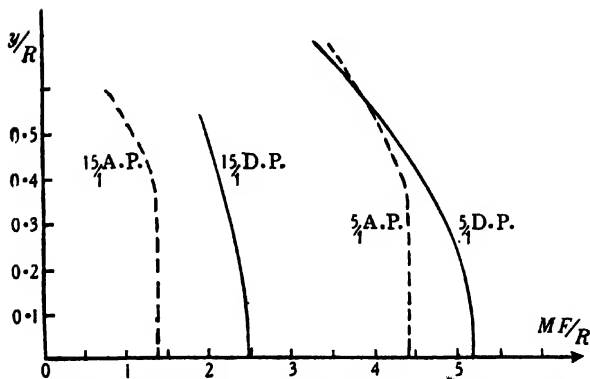


FIG. 67. Spherical aberration curves for two-cylinder lens: y/R = semi-aperture of beam in terms of the tube radius R ; A.P. = accelerating potential; D.P. = decelerating potential. (Klemperer.)

without at the same time causing the negative power of one half of the lens to neutralize the positive power of the other half.

Magnetic lenses suffer from considerable positive spherical aberration, as first observed by Ruska. Recent work by Goddard and Klemperer has shown the aberration curve to have much the same form as that for the decelerating electrostatic lens, there being a negligibly small region near the axis over which the focal length is constant with aperture. Attempts to increase the working aperture of the magnetic objective lens have so far been frustrated mainly by this high value of the spherical aberration.

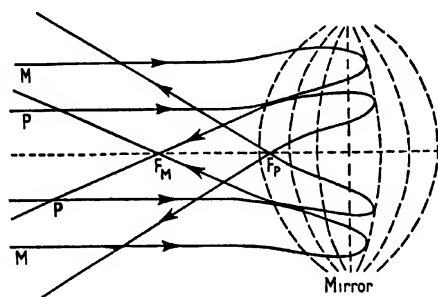


FIG. 68. Spherical aberration of electron mirror.

For other purposes the practical remedy has been to increase the diameter of the coil generating the field, and hence to increase the volume at the centre of the lens over which the field has near-Gaussian characteristics. The sign of the aberration being the same in both electrostatic and magnetic lenses, correction by combination is in principle impossible.

The electron mirror, on the other hand, shows negative spherical aberration when the negative electrode is just below the critical potential. Electrons are then deviated towards the axis, as in a concave mirror, and the marginal paths cross the axis at points farther from the mid-plane than do the paraxial paths (Fig. 68). As in the electrostatic lens, the focusing field is stronger at the margin than near the axis, but the effect on the focus is the reverse since the electrons are doubled back on their tracks by the mirror. In theory, therefore, it is possible to correct spherical aberration by a proper combination of lens and mirror; in practice the difficulties are formidable, as indicated in section 28, on account of the impossibility of extracting the final image from the system without in some way interfering with the focusing fields. Further advance in this direction is largely dependent on the development of the theory and practice of skew rays, where the image is formed on another axis from that containing the object. This branch of the subject, even in light optics, is still in its infancy.

Theoretical Discussion. Progress towards correction of the spherical error has been little greater along theoretical lines. Further development

of the treatment given in the previous section runs into formidable difficulties, and only in special cases have results of practical importance been achieved so far. The simplicity of form of the expression for the spherical aberration (V.21) conceals the complicated form of the coefficient ζ :

$$\zeta = \frac{M}{16r_{\alpha B}^2 \sqrt{V_0}} \int_{z_0}^{z_1} V^{-1} \left[\frac{5}{4} V'^2 + \frac{5}{24} \frac{V'^2}{V^2} + \frac{14}{3} \frac{V' r'_\alpha}{V r_\alpha} - \frac{3}{2} V'^2 \frac{r_\alpha'^2}{r_\alpha^2} + \frac{e}{m} V H'^2 + \right. \\ \left. + \frac{3}{8} \frac{e^2}{m^2} H^4 - \frac{e}{m} V H^2 \frac{r_\alpha'^2}{r_\alpha^2} + \frac{35}{16} \frac{e}{m} \frac{V'^2}{V} H^2 - \frac{3e}{m} V' H H' - \frac{2e}{m} V' H^2 \frac{r_\alpha'}{r_\alpha} \right] r_\alpha^4 dz, \quad (\text{V.22})$$

integrating along the axis from object to image. As before, r_α is a special solution of the paraxial ray equation, M is the magnification relation, and V and H refer to the axial values of the electrostatic potential and the magnetic field respectively.

The evaluation of the definite integral in this expression is a most laborious process. Usually, however, we are concerned with electrostatic or magnetic systems separately, and not in combination. Glaser has shown for the magnetic lens that V.22 may be expressed in a more concise form. Since at the object point $r_\alpha = 0$ and $r'_\alpha = 1$ (cf. p. 118), we have $r_{\alpha B} = z_0$ and $\theta = r_B/z_0 = r_B/r_{\alpha B}$, and the spherical aberration ρ_s is given by

$$\rho_s = \theta^3 C_s \\ = \theta^3 \frac{eM}{96mV_0} \int_{z_0}^{z_1} \left(\frac{2e}{mV} H^4 + 5H'^2 - HH'' \right) r_\alpha^4 dz, \quad (\text{V.23})$$

where θ is the angular aperture of the focused beam. For the special case of a weak lens in which the initial direction of the beam is parallel to the axis ($r_\alpha = \text{const.} = r_{\alpha B} = f$), Scherzer has given an approximate derivation of the radial aberration in the focal plane (ρ_{sf}):

$$\rho_{sf} = -\frac{er_B^3 f}{16mV_0} \int_{-\infty}^{\infty} H_z'^2 dz + \text{terms in } \frac{1}{V_0^2},$$

where f is the focal length for a paraxial ray.

The corresponding expression for the electrostatic lens is

$$\rho_{sf} = -\frac{5r_B^3 f}{64V_0^2} \int_{-\infty}^{\infty} V'^2 dz + \text{terms in } \frac{1}{V_0^3}. \quad (\text{V.24})$$

These formulae allow the aberration to be calculated without undue labour once the form of the field has been determined experimentally.

By repeated partial integration of V.22 Scherzer was also able to obtain it in a form from which it could be shown that the spherical aberration could not be zero in any field which acted as a lens, whether electric, magnetic, or both combined. Absolute correction is therefore unattain-

able, but Rebsch has shown that ρ_s can be reduced below any assigned limit by immersing the object in a strongly varying field, a condition which is equivalent to bringing a supplementary lens close to the object. It has been realized in electron microscopes in which the cathode acts as object, by allowing the field of the objective lens to extend over it, but the problem has not yet been solved for the normal electron microscope, in which an almost parallel beam of electrons from the cathode falls upon an object at some distance from it.

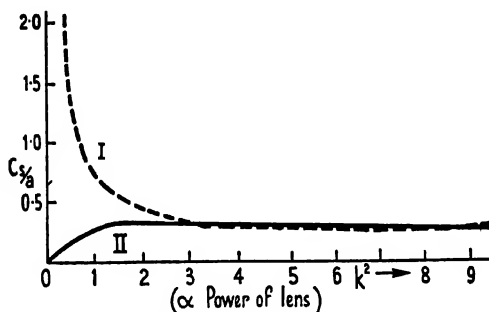


FIG. 69. Spherical aberration of magnetic objective lens (computed). I. Object in position for optimum magnification. II. Object at centre of lens. (Glaser.)

Scherzer attempted to show that the form of field having minimum spherical aberration could be calculated for weak magnetic and electrostatic lenses. Glaser has shown, however, for a special case of the magnetic lens that the value calculated from a strict treatment is below the limit assigned by Scherzer. Taking the type of lens for which the form of the field is very closely reproduced by the function (IV.21),

$$H_z = \frac{H_0}{1 + (z/a)^2},$$

Glaser calculated the spherical aberration factor C_s for various values of the parameter k^2 , from the expression given above (V.23). The relation reduces now to the form

$$\frac{C_s}{a} = \frac{\pi k^2}{4(k^2+1)^{1/2}} \frac{(a^2+z_0^2)^2}{a^4} - \frac{1}{4} \frac{4k^2-3a^2+z_0^2}{4k^2+3a^2+z_0^2} \left[\frac{z_0^2 z_f + a^2(2z_0-z_f)}{a^3} \right], \quad (\text{V.25})$$

where a , z_0 , and z_f have the same significance as before (section 33). The result is shown graphically in Fig. 69, where the upper curve gives the aberration when the object is so positioned in the field as to give high magnification, and the lower curve gives the corresponding values when the object is located at the centre of the lens ($z = z_0$). Minimum aberration is obtained for very weak lenses ($k^2 < 1$) and low magnification, as was to be expected. For values of $k^2 < 1$ attempts to achieve high magnification involve a very high aberration coefficient. High magnification with low spherical error, as is required in the electron microscope,

can only be obtained with strong lenses ($k^2 > 2$). Fortunately the aberration is comparatively small at the minimum focal length which, as previously recorded, occurs for $k^2 = 3$; the upper curve has a very flat minimum at about $k^2 = 6$. The more important quantity in relation to the quality of the image is not C_s/a , however, but C_s/f , as mentioned above. Glaser found that over a wide range of the parameter k this ratio had a value less than the lower limit of 0.25 deduced by Scherzer, and had a minimum value of about 0.17 in the field of special form that he investigated. There may well exist magnetic fields of different axial distribution for which the factor is still smaller.

Although failing in this case, the method of approach adopted by Scherzer may have fruitful results for electrostatic lenses. Taking the case of the weak symmetrical lens, in which the relative change in potential within the lens is small, and using the approximate expression for the spherical aberration (V.24), he proceeded to find the axial field distribution for which the aberration is least. The lengthy argument will not be reproduced here; the conclusion reached is that the required field is of the form

$$V_{0z} = V_0(1 + Ae^{-Bz^2}), \quad (\text{V.26})$$

where A and B are constants (for a given lens) involving the focal length and the separation of the principal planes only. Substituting in expression II.11, we obtain for the distribution throughout the lens:

$$V_{xx} = V_0 + \frac{AV_0}{2\pi} \int_0^{2\pi} e^{-B(x+ir \sin \alpha)^2} d\alpha. \quad (\text{V.27})$$

Scherzer concluded that the electrode form required to produce such a field would be too complicated to construct, being similar to that deduced by Motz for a linearly increasing axial potential (Fig. 15). However, Plass has recently calculated the equipotential distribution given by V.26 for the special case

$$V_{0z} = V_0(1 - \frac{1}{2}e^{-z^2/2}).$$

For small values of the radial distance, r , he found a simple configuration (Fig. 70a), which, in effect, corresponds to the three central elements of the system deduced by Scherzer for large values of r . Since any equipotential surface may be taken as a conducting boundary for the system, it would be entirely possible to construct a lens with electrodes of the form indicated by the outermost curves of the diagram. Plass similarly investigated the two-cylinder lens, using an expression for the axial potential distribution suggested by Gray, in the form

$$V_{0z} = V_0(1 - \frac{1}{2} \tanh z). \quad (\text{V.28})$$

The resulting map of the field is shown in Fig. 70b; as expected on general grounds, the cylinders give less aberration if flared away from each other

as they approach. According to Rebsch and Schneider, however, the potential form for minimum aberration is of a different type:

$$V_s = A \int_0^z e^{-Bs^2} dz,$$

and will result in a slightly different form for the equipotentials.

Ramberg has independently approached the problem in the same way as Glaser and making use of Plass's results. He evaluates the variation in spherical aberration for fields of various refracting powers, and compares the results for different types of lens. The error is in general less for

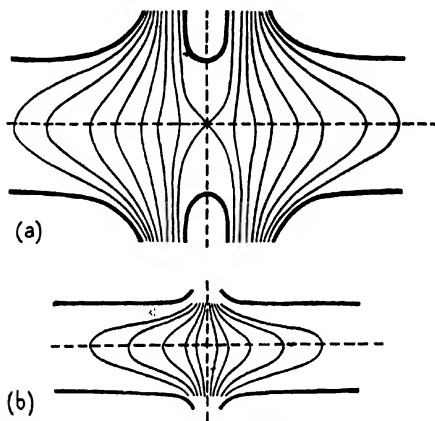


FIG. 70. Electrode forms for minimum spherical aberration (Plass).
(a) Symmetrical lens. (b) Two-cylinder lens.

magnetic than for electrostatic lenses, but is least for the two-cylinder lens, for a given focal length and lens thickness. The variation with focal length has the same form for all lenses, and confirms Glaser's result for a special field (Fig. 69). These results form the only direct aid which the theory of aberrations has yet been able to give in solving the practical problems of electrode design. They have still to be tested and, since they relate to comparatively long focal lengths, it is in any case not certain that they will be applicable to the majority of lens systems.

In electron optical practice, and especially in the electron microscope, 'correction' of spherical aberration has been obtained by the expedient of increasing the accelerating voltage applied to the beam. As V.22 shows, the error is roughly inversely proportional to the square of the voltage in both types of lens, and it is therefore beneficial to employ the highest anode potential compatible with other experimental conditions. At the same time, the use of higher voltages results in greater brilliance of the image thrown on the fluorescent viewing screen, and consequently smaller apertures may be inserted, resulting in further reduction in spherical

aberration. In the electron microscope the working aperture is of the order of 10^{-3} radian only; indeed it is often so small that diffraction effects have to be taken into account in estimating the possible resolution of the instrument. Progress towards higher resolving powers, of the order of atomic dimensions, will depend on improvement in design of the pole-pieces generating the field, so that the region over which spherical aberration is negligible is considerably increased from its present small value. A further possibility of correction, by the controlled use of space charge within the lens, is discussed in section 44.

*40. Distortion

The first term, αr_a^3 , in the total aberration equation (V.20) contains only the radial coordinate at the origin, and is independent of the position of a ray in the aperture. Hence the value of this error will be the same for all rays from a given object point, no matter what their direction with respect to the axis. As a result there will be no lack of sharpness of the image points in the Gaussian image plane, but they will be shifted in a uniform manner from the positions predicted by the first-order theory; there will be distortion of the image, but no confusion. The form of the error which occurs both in light optics and in electrostatic lenses is equivalent to a variation of magnification with radius: the outer parts of the image are enlarged or diminished with respect to the axial region. The image of a regular graticule would appear distorted in the 'barrel' or 'pincushion' manner, as shown in Figs. 71a and b respectively. The aberration coefficient is given by the definite integral

$$\begin{aligned} \alpha = & \frac{M}{16\sqrt{V_0}} \int_{z_0}^{z_1} V^{-1} \left[\frac{5}{4} V'^2 + \frac{5}{24} \frac{V'^4}{V^2} + \frac{7}{2} \frac{V'^2}{V} \left(\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{3r_\alpha} \right) - \frac{3}{4} V'^2 \left(\frac{r_\gamma'^2}{r_\gamma^2} + \frac{r_\alpha' r'_\alpha}{r_\gamma r_\alpha} \right) + \right. \\ & + \frac{e}{m} V H'^2 + \frac{3e^2}{8m^2} H^4 - \frac{eV}{2m} H^2 \left(\frac{r_\gamma'^2}{r_\gamma^2} + \frac{r_\gamma' r'_\alpha}{r_\gamma r_\alpha} \right) + \frac{35e}{16m} \frac{V'^2}{V} H^2 - \\ & \left. - \frac{3e}{m} V' H H' - \frac{eV'}{2m} H^2 \left(\frac{3r_\gamma'}{r_\gamma} + \frac{r'_\alpha}{r_\alpha} \right) \right] r_\gamma^2 r_\alpha dz + \\ & + \frac{M}{4} \left[\left(\frac{V''}{8V} + \frac{5V'^2}{32V^3} + \frac{V' r'_\gamma}{2V r_\gamma} + \frac{3}{2} \frac{r_\gamma'^2}{r_\gamma^2} + \frac{eH^2}{8mV} \right) r_\gamma^2 \right]_{z_0}^{z_1} + \\ & + \frac{M}{16} \sqrt{\left(\frac{2e}{m} \right)} \int_{z_0}^{z_1} V^{-1} \left[V^2 r'_\gamma - \frac{V'}{2} r_\gamma r'_\gamma + r_\gamma^2 \left(\frac{3eH^2}{8m} + \frac{9V'^2}{8V} - V' \frac{H'}{H} \right) \right] H dz + \\ & + \frac{M}{32} \sqrt{\left(\frac{2e}{m} \right)} \left[V^{-1} \left(\frac{3}{2} \frac{V'}{V} + \frac{2r'_\gamma}{r_\gamma} - \frac{H'}{H} \right) H r_\gamma^2 \right]_{z_0}^{z_1} \quad (\text{V.29}) \end{aligned}$$

containing, in the general case of combined fields, an imaginary part due to the magnetic field as well as a real part corresponding to the normal optical defect. The occurrence of barrel or of pincushion distortion

depends on whether the real part is negative or positive. The imaginary part relates to the rotation of the rays in a magnetic field, and indicates a rotation of the outer parts of the image with respect to the axis. When

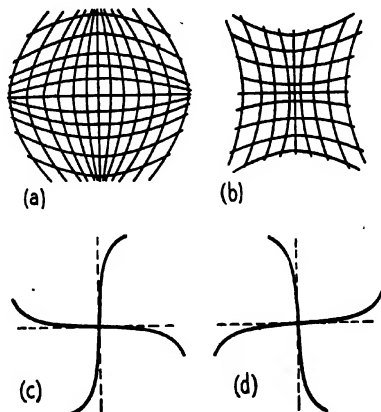


FIG. 71. Distortion in electron lenses: (a) barrel type; (b) pincushion type; (c) rotational effect in magnetic field (coefficient negative); (d) rotational effect in magnetic field (coefficient positive).

this part of the term is negative, the rotation is clockwise, and counter-clockwise when the value is positive, as shown in Figs. 71c and d respectively.

Electrostatic lenses appear to show mainly distortion of the pincushion type, whilst electron mirrors may produce either type, but its magnitude is never large in practical electrode systems; frequently, as in the electron microscope, when the clarity of the image is the first requisite and only a very small portion of the object is imaged, it is in any case not of great importance.

When magnetic lenses are used alone, the spiral error can be corrected by using two semi-lenses with coils wound in opposite directions, in place of the normal type. The equal and oppositely directed fields cancel each other in respect of rotation of the image, which depends on the integral of the field strength along the axis, whereas their focusing effect is additive, being dependent on the square of the field. For the general case of combined electric and magnetic fields it is possible to design a lens free from distortion by making the electric field symmetrical, the magnetic field unsymmetrical, about the mid-plane of the lens. It is not usually necessary to resort to such special measures, however, since the limitations introduced in order to reduce spherical and chromatic aberration at the same time largely eliminate the smaller distortion defect.

*41. *Curvature of the Field and Astigmatism*

The type of aberration represented by the second term of V.20, (β), is very different from that given by the first and last, but closely related to

that of the following term, (γ) . If AC is the (principal) ray from a point object A through the centre C of the aperture of a lens (Fig. 72), suppose a number of other rays from the same point to pass the aperture all at the same distance $|u_B|$ from C , as represented by a circle drawn with this radius. Then, in the presence of the second aberration only, this bundle of rays will pass the Gaussian image plane in a circle of radius $r_i (= \beta r_a^2 u_B)$

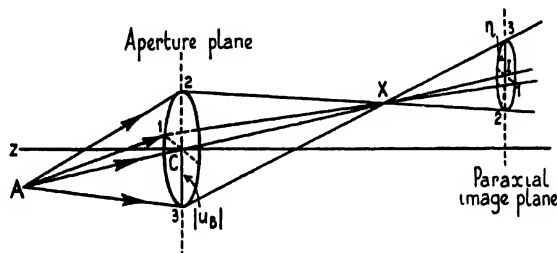


Fig. 72. Curvature of the field: an off-axis point A gives a point image at X , but a circle of confusion in the paraxial image plane.

about a centre on the principal ray A , intersecting each other in a common point X which is usually before the image plane, as shown. There will thus be no clear image formed in the plane I , but an undistorted point image will be formed at X . The position of X will vary with the position of the object point; when the latter is confined to a plane perpendicular to the axis, the X -surface will be spherical and concave towards the object. It may be shown from V.30 that its radius of curvature, R_x , is related to the aberration coefficient β by the expression

$$1/R_x = 2\beta \frac{r_{\alpha B}}{M} \sqrt{\frac{V_i}{V_0}}.$$

The value of the coefficient β is given by the definite integral:

$$\begin{aligned} \beta = & \frac{M}{8r_{\alpha B}\sqrt{V_0}} \int_{s_0}^{s_i} V^{-1} \left[\frac{5}{4} V'^2 + \frac{5}{24} \frac{V'^4}{V^2} + \frac{7}{3} \frac{V'^3}{V} \left(\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{r_\alpha} \right) + \right. \\ & + V'^2 \left(\frac{r'_\gamma}{4r_\gamma^3} + \frac{r'_\alpha}{4r_\alpha^3} - \frac{2r'_\gamma r'_\alpha}{r_\gamma r_\alpha} \right) + \frac{e}{m} V H'^2 + \frac{3e^2}{8m^2} H^4 + \frac{e}{m} V H^2 \left(\frac{r'_\gamma}{2r_\gamma^3} + \frac{r'_\alpha}{2r_\alpha^3} - \frac{2r'_\gamma r'_\alpha}{r_\gamma r_\alpha} \right) + \\ & + \frac{35e}{16m} \frac{V'^2}{V} H^2 - \frac{3e}{m} V' H H' - \frac{e}{m} V' H^2 \left(\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{r_\alpha} \right) \left. \right] r_\gamma^2 r_\alpha^2 dz + \\ & + \frac{M}{8r_{\alpha B}} [\sqrt{V_0} V' V' + 4r'_\gamma r'_\alpha]_{s_0}^{s_i}. \end{aligned} \quad (V.30)$$

Curvature of the field is usually found associated with astigmatism of the image, which results in two line foci appearing for each cone of rays from an object point off the axis, and, for an extended object, in the separation of each of the image surfaces discussed above into two surfaces of

differing curvature. The third term, γ , in the aberration equation (V.20) relates to astigmatism. It will be seen that it differs in form from the β -term only in that it contains the conjugated form (\bar{u}_B) of the paraxial ray solution. This difference has the physical meaning (in terms of Fig. 72 above) that a series of rays designated 1, 2, 3,... clockwise around the circle of radius $|u_B|$ in the aperture, are found to pass the circumference of a circle in a second plane in counter-clockwise order. It follows that

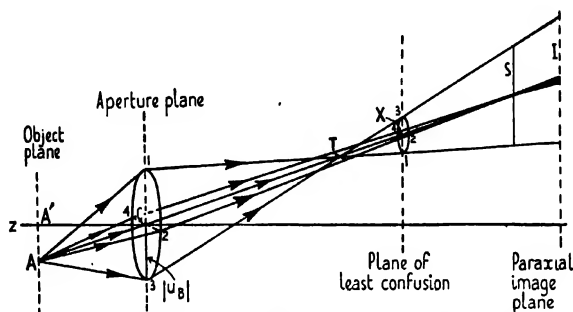


FIG. 73. Astigmatism: an off-axis point A gives two line foci, the tangential (T) and sagittal (S).

corresponding points on the two circles must be displaced by a varying angle about the direction of the principal ray A ; in fact only two pairs of rays remain throughout in a plane containing the latter. This rather complex set of paths is indicated in Fig. 73, where the four special rays are shown numbered. The rays 1 and 3 are in the (meridional) plane containing the axis and the object, here considered as a line perpendicular to the axis. The rays 2 and 4 are in a plane passing through the axis at right angles to the meridional plane; this is termed the sagittal plane.

For a point on the axis the focal length of a lens of rotational symmetry is necessarily the same in all planes through the axis. But for a point off the axis, such as A , this is no longer the case. It can be shown that the focal length in the sagittal plane is slightly less than the paraxial value; in the glass-optical sense, this is due to the slightly larger angle of incidence at the lens surface of rays such as 2 and 4, compared with a ray to the same points from the axis (A'). In the meridional plane the focal length is less still, mainly on account of the extreme obliquity of the limiting ray 1. The result is that rays 1 and 3 are brought to a focus at a point T , the tangential focus, and rays 2 and 4 to a more distant point, the sagittal focus S , both on the principal ray ACI . As already mentioned, these are the only rays which intersect the principal ray ACI after refraction. Detailed ray-tracing shows that all other rays pass through two line images, one at T , perpendicular to the meridional plane, the other at S , perpendicular to the sagittal plane and therefore in the meridional plane; the length of each line is determined, for a given lens, by the radius of the circular zone involved. The appearance of the corresponding image in the

Gaussian image plane I will be that of an ellipse, with major and minor axes determined by rays 2 and 4, and 1 and 3 respectively. The circle at X , between T and S , constitutes a circle of least confusion.

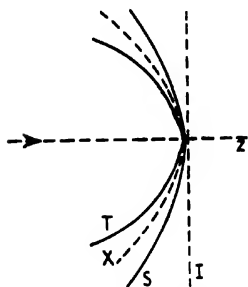


FIG. 74. Image surfaces in astigmatic lens: I = paraxial image plane, S = sagittal image surface, T = tangential image surface, X = surface of least confusion.

When the object is not a point but a surface perpendicular to the axis at A' , then the image lines T and S will fill two surfaces both concave towards the lens, but elliptical instead of spherical in form; they are tangential to each other and to the Gaussian image plane (I) at the axis (Fig. 74). The disk of least confusion will likewise become a concave surface (X), intermediate between the other two, and will now represent the surface upon which the clearest image may be found. It may be shown, from V.31 below, that the radii at the axis of the S and T surfaces are given in the electron optical case by

$$R_S - R_T = 4|\gamma| \frac{r_{\alpha B}}{M} \sqrt{\frac{V_i}{V_0}},$$

where V_0 , V_i , and M have the same meaning as before and γ is the value of the astigmatism coefficient. There is an increase in the size of the disks of least confusion with radius in the intermediate plane, the definition decreasing rapidly towards the outer edge of the field. In the normally observed case, γ is positive and the tangential surface has greater curvature than the sagittal surface.

The value of γ is given by the definite integral:

$$\begin{aligned} \gamma = & \frac{M}{16r_{\alpha B}\sqrt{V_0}} \int_{z_0}^{z_i} V^{-1} \left[\frac{5}{4} V'^2 + \frac{5}{24} \frac{V'^4}{V^2} + \frac{7}{3} \frac{V'^3}{V} \left(\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{r_\alpha} \right) - \right. \\ & - V'^2 \left(\frac{5}{4} \frac{r_\gamma'^2}{r_\gamma^2} + \frac{5}{4} \frac{r_\alpha'^2}{r_\alpha^2} - \frac{r'_\alpha r'_\gamma}{r_\alpha r_\gamma} \right) + \frac{e}{m} V H'^2 + \frac{3e^2}{8m^2} H^4 - \frac{e}{m} V H^2 \left(\frac{3r_\gamma'^2}{2r_\gamma^2} + \frac{3r_\alpha'^2}{2r_\alpha^2} - \frac{2r'_\gamma r'_\alpha}{r_\gamma r_\alpha} \right) + \\ & + \frac{35e}{16m} \frac{V'^2}{V} H^2 - \frac{3e}{m} V' H H' - \frac{e}{m} V' H^2 \left(\frac{r'_\gamma}{r_\gamma} + \frac{r'_\alpha}{r_\alpha} \right) \left. \right] r_\gamma^2 r_\alpha^2 dz + \frac{M}{4r_{\alpha B}} [r'_\gamma - r'_\alpha]_{z_0}^{z_i} + \\ & + i \frac{M}{8r_{\alpha B}\sqrt{V_0}} \left(\frac{2e}{m} \right) \int_{z_0}^{z_i} V^{-1} \left[V r'_\gamma r'_\alpha - \frac{V'}{4} (r_\gamma r'_\alpha + r_\alpha r'_\gamma) + \right. \\ & \left. + r_\gamma r_\alpha \left(\frac{3eH^2}{8m} + \frac{9V'^2}{8V} - V' \frac{H'}{H} \right) \right] H dz + \\ & + i \frac{M\sqrt{V_0}}{16r_{\alpha B}\sqrt{V_0}} \left(\frac{2e}{m} \right) \left[\frac{H}{V} \right]_{z_0}^{z_i}. \end{aligned} \quad (\text{V.31})$$

The imaginary portion characterizes a rotational effect produced by the magnetic field H and termed 'anisotropic astigmatism'. It results in the

tangential and sagittal focal lines being rotated from the positions indicated above; they still remain at right angles to each other, however.

The terms 'tangential' and 'sagittal' derive from consideration of a wheel object. Any point on the rim will produce a line image T parallel to the tangent to the wheel at the object point; hence the term 'tangential focus'. For the whole circumference of the wheel these T -lines will form a continuous circular image; in other words, the rim of the wheel will be sharply in focus on a screen intersecting the T -plane at the correct distance from the axis (depending on the magnification). In the same way the spokes will throw a series of line images S beyond the T -lines and at right angles to them; they will thus be imaged as a set of radial lines in the curved S -plane. On withdrawing the screen from the position in which the rim was in sharp focus the outer ends of the spokes will first appear on it as short radial lines directed towards the axis. Moving the screen farther back will bring the other parts of the spokes into focus in turn; they will always appear as a set of short lines pointing towards the axis, as arrows might. Hence the origin of the term 'sagittal' to denote this focal surface. Finally, the hub of the wheel will appear in sharp focus when the screen reaches the paraxial image point.

In general, both curvature of the field in the simple sense and astigmatism are present together, with the result that the tangential and sagittal surfaces (and that of least confusion) are curved to a greater extent than calculated for astigmatic aberration alone. In principle both may be corrected by application of the Petzval condition to the construction of a lens combination:

$$N_1 f_1 = N_2 f_2,$$

where N_1 and f_1 refer to the refractive index and focal length of the first component and N_2 and f_2 to the corresponding values for the second component. This condition holds true for any separation of the lenses and is the basis for correcting the two defects in light optics. Klemperer and Wright have discussed the possibility of correcting an electrostatic two-cylinder lens along similar lines, regarding the field as divided up into a large number of segmental lenses. Unfortunately it appears that the effect of the positive components will always predominate, giving an image surface strongly concave towards the lens, so that direct correction of the isolated lens would seem to be impossible. Voit showed, however, that, with the aperture in a particular position and a fixed magnification of 6.16, there is one form of field distribution for which astigmatism vanishes.

In most electron optical devices the size of the object and of the aperture are both small and the curvature errors do not obtrude themselves. The most notable exception is the image tube, as described in section 27 (p. 70). By the nature of the case the object is here large and the image is required to be of a similar size. As both β - and γ -terms involve the square of the radial distance of the object point; they will have a great effect on the outer parts of the surface to be reproduced. Morton and Ramberg made a thorough investigation of this type of tube and found

that the radius of curvature of the T -surface (at the axis) was as small as $0.044R$ and that of the S -surface $0.075R$, where R is the radius of the lens tube. The practical remedy is found in curving the cathode towards the lens, so that the curvature of the image surface is correspondingly reduced. The effect of the aperture size (u_B) is largely eliminated by employing high potentials to accelerate the emitted electrons and thus confine them to very narrow pencils. The definition of the resulting image is remarkably good in both the Zworykin and Farnsworth types of image transformer.

*42. *Coma*

The remaining two terms (δ and ϵ) of the aberration equation relate to another complex phenomenon, coma, in which beams through different zones of the lens form image-disks of varying diameter, with centres displaced along a line. It is due to the asymmetrical position of the object, as is astigmatism, and is serious even for points close to the axis of the lens. Both terms are dependent on the radial ordinate of the object point and on the square of the ordinate of the ray in the aperture. The δ -term determines the size of the image-disk and the ϵ -term the displacement of its centre from the true image point; they are not independent, being conjugately related.

If astigmatism is absent, then any beam from a given object point will form a point image; this is the definition of a stigmatic beam. If spherical aberration also is absent, all beams through different zones of the lens will produce such point images in the true image plane, which will be flat when the curvature term is zero. The remaining possible aberration is that the image points will fall in different parts of this plane, and in fact they are found to produce an extended image which bears a marked resemblance to the bright head and flared tail of a comet; whence the name of coma. The total effect on the image of an extended object is one of greater confusion than that caused by any other single aberration.

Considering first the rays in a meridian plane from an off-axis point A (Fig. 75a), we suppose the paraxial image to be the point I in which the principal ray through the centre (B) of the aperture meets the image plane. Then the term in ϵ corresponds to a displacement of the image point towards or away from the axis along the normal from I , for rays of increasing aperture. Typical paths are shown in the diagram; when the sign of the ϵ -term is negative, the direction of aberration is towards the axis. The effect can be interpreted as a variation of magnification for different zones in the lens as compared with the paraxial value. For skew rays, however, having paths out of the plane of the diagram, there is a displacement in the x - as well as in the y -direction, the amount of which depends on the value of the x -coordinate of the ray in the aperture. Considering a given zone (Fig. 75c), we find that pairs of rays incident at opposite ends of a diameter are focused at the same point, but that as this diameter is rotated about the centre B the image point describes a circle in the same sense but at twice the angular speed. Thus the meridian rays

are focused at a point 1, nearest to I , the pairs displaced 45° from it at points 2 and 4 on a line at right angles to II' , and the sagittal rays (in the astigmatic sense) are focused on the same line as the point 1, but farther from the Gaussian image point. A cone of rays from this zone is therefore focused as a comatic circle of radius determined by the coefficient δ , with its centre displaced from I by an amount determined by ϵ . When all the

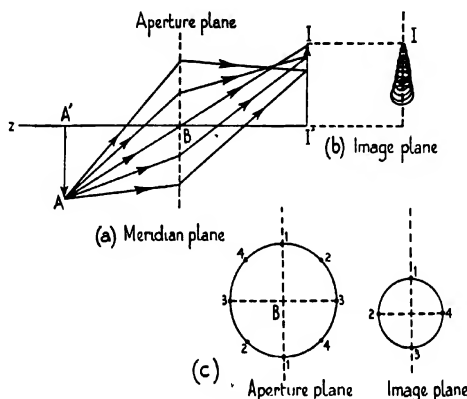


FIG. 75. Origin of coma.

zones of the lens are considered, the total effect on the image is that shown in Fig. 75: a bright point focus with a broadening tail, the intensity of which decreases rapidly with distance from I .

In electron lenses the tail of the image is directed away from the axis, in the opposite sense from that shown in Fig. 75b. As the defect occurs for points close to the axis it is necessary to ensure strict alinement of all electrodes on the axis of the system, even when beams of small aperture are in use. When beams of larger aperture are employed the true coma is usually not great if this precaution has been taken. For complete correction the Abbe condition must be satisfied:

$$\sin \theta_1 / \sin \theta_2 = \text{constant},$$

where θ_1 is the angle made with the axis by the incident beam and θ_2 that made by the emergent beam. In the general case, when there exist media of differing refractive index N_1 , N_2 on either side of the system, the rule takes the form of a revision of Lagrange's law (III.5),

$$r_a N_1 \sin \theta_1 = r_i N_2 \sin \theta_2,$$

where r_a and r_i represent the radial distances of object and image point respectively. This is equivalent to the statement that the lateral magnification will be constant for all zones if the form of the field is such that $\sin \theta_1 / \sin \theta_2$ is constant, and coma cannot then appear. Under such conditions spherical aberration cannot be present either; in optics such a lens

is said to be aplanatic. As stated in section 39, it has been established that full correction of spherical aberration is impossible in electron lenses, and hence no truly aplanatic lens can exist. It follows, however, that in devising conditions in which spherical aberration is reduced to a minimum, the (quantitatively less) comatic error is also reduced; it is, indeed, rarely visible in practical electron lenses.

The coefficients δ and ϵ are given by the definite integral

$$\begin{aligned} \epsilon = 2\bar{\delta} = & \frac{M}{8r_{\alpha B}^2 \sqrt{V_0}} \int_{z_0}^{z_i} V^{-1} \left[\frac{5}{4} V'^2 + \frac{5}{24} \frac{V'^4}{V^2} + \frac{7}{2} \frac{V'^3}{V} \left(\frac{r'_\gamma}{3r_\gamma} + \frac{r'_\alpha}{r_\alpha} \right) - \frac{3}{4} V'^2 \left(\frac{r'_\gamma r'_\alpha}{r_\gamma r_\alpha} + \frac{r_\alpha'^2}{r_\alpha^2} \right) \right. \\ & + \frac{e}{m} V H'^2 + \frac{3e^2}{8m^2} H^4 - \frac{eV}{2m} H^2 \left(\frac{r'_\gamma r'_\alpha}{r_\gamma r_\alpha} + \frac{r_\alpha'^2}{r_\alpha^2} \right) + \frac{35e}{16m} \frac{V'^2}{V} H^2 - \\ & \left. - \frac{3e}{m} V' H H' - \frac{eV'}{2m} H^2 \left(\frac{r'_\gamma}{r_\gamma} + \frac{3r'_\alpha}{r_\alpha} \right) \right] r_\gamma r_\alpha^2 dz + \frac{M}{4r_{\alpha B}^2} [r_\alpha'^2]_{z_0}^{z_i} + \\ & + i \frac{M}{8r_{\alpha B}^2} \sqrt{\frac{2e}{m}} \int_{z_0}^{z_i} V^{-1} \left[V r_\alpha'^2 - \frac{V'}{2} r_\alpha r'_\alpha + r_\alpha^2 \left(\frac{3eH^2}{8m} + \frac{9V'^2}{8V} - V' \frac{H'}{H} \right) \right] H dz, \end{aligned} \quad (V.32)$$

with the symbols bearing the meanings already assigned. It will be seen that it contains an imaginary as well as a real part, the latter relating to the isotropic effect discussed above, and the former corresponding to a rotational effect produced by the magnetic field, as in the case of distortion and astigmatism. This anisotropic coma is observable as a rotation of the image tails about the paraxial image point, so that they no longer are directed along the line joining it to the axial image point. As in the case of anisotropic astigmatism, the rotation can be corrected by using a lens composed of two oppositely directed magnetic fields, which produce equal and opposite rotations of the rays but are additive in their focusing effect. If one of the anisotropic errors is corrected in this way, the remaining two are automatically eliminated also. The whole of the comatic term, as well as the distortion term, can in principle be reduced to zero by the formation of a special type of combined electrostatic and magnetic field, but the small proportion of the error makes it unnecessary to resort to this means of correction. In image tubes of the Zworykin or Farnsworth type coma is effectively eliminated by the use of high accelerating potentials, which limit the beam of electrons from a given image point to a very small aperture.

*43. Chromatic Aberration

In the above discussion of third-order aberrations we have assumed the incident electron beam to be uniform in velocity. Actual electron beams, however, will exhibit a certain distribution or 'straggle' of velocities in any cross-section, or with time in any single ray-path. The cause resides in the inevitably finite velocity of emission of electrons from the cathode (of whatever type) and may be aggravated by fluctuations in the accelerat-

ing potential applied to it. Electrons are then refracted by a lens field in differing degree according to their initial velocities. If the incident beam is parallel to the axis, then in a positive lens the slower electrons will suffer greater deviation than the faster, and will pass through a focus nearer to the lens. Even when all other aberrations are absent, therefore, a series of foci will result (Fig. 76) and the axial distance between that of the fastest (F_f) and slowest (F_s) electrons will depend on the range of velocities and on the form of the focusing field. The degree of aberration

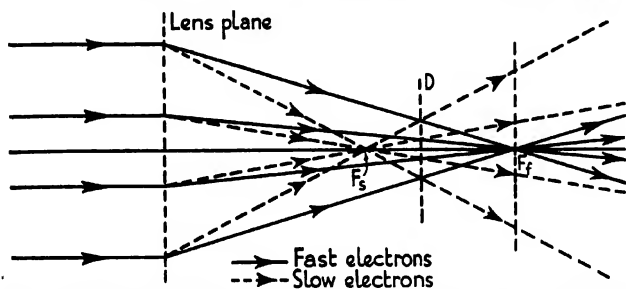


FIG. 76. Chromatic aberration of initially parallel pencil. (D = disk of least confusion.)

may be measured longitudinally in terms of F_f, F_s , or as the radius of the disk of least confusion, or by the radius of the disk of confusion in the focal plane corresponding to a particular velocity of the beam. For purposes of practical analysis it is convenient to treat the applied accelerating voltage (V_0) as the reference quantity, and to measure the aberration as the size of the disk of least confusion or as the difference in focal lengths.

The similarity to spherical aberration is obvious, but it must be emphasized that the chromatic error is entirely independent of all others, and will exist even in a thin lens under paraxial conditions, as it is a property of the focusing medium *per se* and not of the size of the aperture or the radial distance of the object point. Off-axial points will suffer in addition a lateral aberration in the $x-y$ plane and, in the presence of magnetic fields, a difference in rotation about the axis, since the expression for ϕ (IV. 13) shows that it depends on the initial velocity of the electron. Here, however, we shall confine ourselves to consideration of the axial aberration.

It follows from the nature of the focusing action of electric and magnetic fields that the chromatic error is unavoidable in all types of lens. If a lens or mirror is to deviate a beam at all, the form of the expressions previously obtained for the focal length shows it to depend on the velocity of the electrons. Hence any initially parallel unhomogeneous beam must emerge from the system as a chromatically separated bundle of rays. All the lenses treated above, including the weakly negative aperture lens, have a focal length that may be approximately represented by

$$f = KV_0,$$

where K is a constant for a given electrode system or magnetic shield, as the case may be. If, now, the equivalent voltage of thermionic emission of the electrons is V_e , we must write

$$f = K(V_0 + V_e).$$

The variation in focal length with emission voltage is given at once as

$$df/dV_e = K = f/V_0,$$

for a given accelerating potential V_0 , and the relative chromatic aberration (longitudinally) as

$$df/f = dV_e/V_0. \quad (V.33)$$

Since it is impossible to decrease the velocity of ejection of the electrons to zero, it follows that the chromatic error will always be finite; and since all electron lenses are positive, with the unimportant exception of the negative aperture lens, the error will always be positive and fast electrons will be focused farther from the lens than slower members of the beam. Correction of a lens for chromatism is therefore impossible by the formation of a suitable combination. In the electron mirror, on the other hand, the error is of opposite sign since the slower the electron the less deeply does it penetrate into the field before reflection occurs, with the result that the focal length is greater for slow than for fast electrons. In principle, therefore, the mirror might be used to correct the chromatic, as well as the spherical aberration of lenses; but, for reasons discussed earlier, the difficulties in the way of realizing such a combination have not yet been surmounted.

The remedy applied in practice has been that immediately suggested by expression V.33: to keep the emission velocity as low and the accelerating potential as high as possible. The first requirement is met by employing 'dull-emitting' oxide-coated cathodes, so that the velocity of thermal emission is small. In many types of electronic apparatus the value of the anode potential is so high that the chromatic error is in any case negligible. Where purity of the final image is a first requirement, as in the electron microscope, the potential is raised to the highest value that can conveniently be employed: in special models voltages as high as 300 kV. have been applied. It is clearly necessary also to stabilize the voltage supply at a similar level of accuracy, so that the maximum fluctuation ΔV_0 is of the same order as (or less than) the value of V_0 . In the case of the cathode-ray tube and similar medium-voltage devices, this is readily achieved by careful attention to the conventional rectifying system, but in the electron microscope the required stability is approximately 1 part in 10^4 . The problem can then only be solved by special adaptation of modern radio-circuit technique to the high-voltage conditions involved; by the use of resonant circuits and feed-back this degree of stability can be reached, as explained in section 61 (p. 209).

When the incident electrons are not a parallel beam, but diverge from a point object, the chromatic defect will appear as a variation in image

distance with velocity and hence its degree will be dependent on the position of the object. This becomes an important consideration in the electron microscope, where the highest possible correction is demanded. Glaser has investigated it for the magnetic immersion objectives previously described, in which the field approximates to a simple form (IV. 21). He finds an expression for the radius of the disk of confusion (r_c) in the image

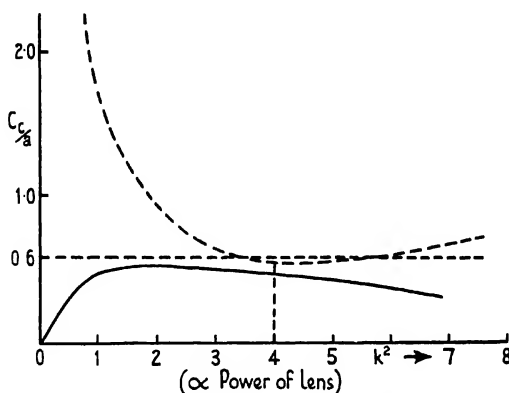


FIG. 77. Chromatic aberration of magnetic lens (computed). I. Object in position for optimum magnification. II. Object in position for least aberration. (Glaser.)

plane corresponding to the applied voltage V_0 , when a very small voltage fluctuation ΔV is present. If M is again the magnification and θ is the angular aperture of the beam proceeding from the object, then

$$r_c = \theta M \frac{\Delta V}{V} \int_{z_0}^{z_i} r_\alpha'^2 dz = \theta M \frac{\Delta V}{V} C_c, \quad (\text{V.34})$$

where r_α is a special solution of the paraxial ray equation such that $r_0 = 0$ and $r_0' = 1$, and C_c is a constant determined by the characteristics of the lens. In the case investigated the value of the constant proves to be

$$C_c/a = \frac{\pi k^2}{2(k^2+1)^{\frac{1}{2}}} \frac{a^2+z_0^2}{a^2}, \quad (\text{V.35})$$

where z_0 is the axial position of the object and a and k have the meanings previously assigned; this expression should be compared with that obtained (V.25) for the spherical aberration of the same lens. The variation of C_c with the power of the lens and the position of the object is indicated by the curves of Fig. 77. The full curve corresponds to the location of the object in the mid-plane of the lens ($z_0 = 0$), and has a maximum for a moderately strong lens (at $k^2 = 2$). The broken line is obtained when the object is placed in the position required for high magnification ($z_0 \approx z_f$, the mid-focal length); in this case the error is of large value in weak lenses and has a flat minimum for $k^2 = 4$. Hence, as

in the case of spherical aberration also (cf. Fig. 69), the chromatic error is small in the strong lenses used in microscopy, for which $k^2 \approx 3$. It will be seen that the value of C_c is higher than that of C_s , except in weak lenses, but since the spherical error depends on the third power of the aperture, whereas the chromatic error is directly proportional to it, the definition of the final image is limited more by the former than the latter. In any case the circuit problems involved in minimizing the chromatic effect prove to be more tractable than the discovery of fields of low spherical aberration.

Ramberg has independently evaluated the chromatic error for electrostatic fields of simple form as well as for the magnetic objective. He finds that the latter exhibit roughly one-half the aberration of electrostatic lenses when the fields are weak, and that the ratio is 1 to 5 or 6 in favour of the magnetic type when the focal length is short.

*44. *Correction by Space Charge; the Ideal Lens*

The Problem of Correction. Summing up the discussion of third-order aberrations we may say that up to the present the problem of correction has been avoided rather than solved. In the majority of electron optical devices the possible defects are largely eliminated by the use of small apertures and high voltages. Where large cathodes are an essential feature of the apparatus, as in the television image tube, the effective aperture may still be kept small by applying a high potential in order to accelerate the beam from each object point into a narrow pencil, and the defects occasioned by extended objects may be minimized by curving the photocathode; the optical designer then has to solve the problem of throwing a similarly curved light image on to the cathode surface.

The progress made towards positive correction of errors is so far very small; neither the empirical approach of modifying the electrode form and studying the result, nor fundamental investigation of the dependence of the degree of aberration on the field distribution, has yet led to major improvement. The principle enunciated by Rebsch, of immersing the object in a region where the rate of change of field strength is high, has been applied successfully to some special cases, but not yet to those lenses of high power where correction is most needed. The more general conclusions of Voit are even more difficult of application. The theoretical possibility of aberrationless combinations of mirrors and lenses has failed of realization to an even greater extent, in face of the severe difficulties involved. It is not too much to say that the practical men have been inclined to take a short cut to a working solution suitable to the degree of refinement at present demanded of electron optical systems, and that the theoreticians, however successful in analysis, have not provided an indication of the way to a practicable solution. It is not out of place, however, to remember that the correction of light optical lenses has engaged the attention of scientists for very many decades whereas electron optics

as a special subject is not yet twenty years old; progress will have been rapid if we approach the perfection of glass lenses in the next two decades.

There is an evident need for intensive investigation of the aberrations of electronic systems with a higher degree of accuracy than has yet been attempted, both by direct experiment and by the methods of ray-tracing long used by optical designers. On the other hand, it is necessary to beware of following too closely the lines along which correction has been achieved in light optics. As previously emphasized, there are considerable and fundamental differences between the two types of system and the success with which some methods of treatment can be applied to both should not blind us to the need for seeking new methods specific to the electronic problem. In particular the theoretical discussion of systems of minimum aberration is much simpler in the case of electromagnetic fields of continuously variable strength (or refractive index). The preoccupation with boundary surfaces of a special form, which is hardly avoidable in glass optics on account of the difficulty of preparing other than spherical surfaces, has invaded electron optics equally, although very wide possibilities of variation of the field are here available. The direction in which progress must be sought is probably that of the investigation and calculation of special electrode forms for which one or more of the different aberrations is a minimum. An important step on this road would be made if the total aberration equation (V.17) could be expressed in a more tractable and physically meaningful form.

The Effect of Space Charge. There remains, however, one possibility of correcting defects which has not yet received mention: the controlled use of space charge. Consideration of electron lens aberrations shows that they are caused by the parts of the field farther from the axis being stronger than required to fulfil the laws of Gaussian optics. Reference to the paths of typical rays in the case of spherical or chromatic aberration (Figs. 64 and 76) will show that correction would be achieved if the field were slightly weakened to an extent depending on the radial distance from the axis. If the form of the field be regarded as given by the form and disposition of the electrodes, then the only means of changing it is by introducing charged particles into the lens space; in previous discussion we have always assumed the system to be free of space charge. The general form of the charge distribution required for correction follows at once from the condition that paraxial rays must be given slightly more deviation and marginal rays slightly less. Such an effect would be produced by establishing a positive space charge at the axis, decreasing in density with radius and possibly becoming negative in the outer zones. This qualitative prediction has been substantiated by Gabor, who succeeded in obtaining an expression for the required distribution of density, and hence the form of the perfect lens.

Applying Fermat's optical principle of least time to the electron, we may state that the path followed by an electron between two points must be such that the time is a minimum, compared with other paths joining

the points. Its application to the motion of a corpuscle was first made by Hamilton over a century ago and forms the basis of his method of treating optical systems, as will be further explained in the Appendix (p. 249). Here it will suffice to state it in the form

$$\int dt = \int ds/u = (1/c) \int N ds \rightarrow \text{minimum}, \quad (\text{V.36})$$

where u is the velocity of the electron, c that of light, N the refractive index of the medium, and the integration is to be taken along the whole path, of which ds is an element. Gabor proceeds to write

$$N ds = dS, \quad (\text{V.37})$$

where dS is the corresponding element of length in a system of uniform refractive index.

If the distribution of potential, and therefore of refractive index, is known for the meridian plane of an axially symmetrical field, we may now transform it into a new surface in which every line element about a point has a value dS , related to that of the corresponding line element ds at the equivalent point in the median plane by expression V.37. For a trajectory in the new surface thus mapped, the Hamiltonian criterion becomes

$$\int dS \rightarrow \text{minimum},$$

so that it is now the shortest distance across the surface from object to image point. The electron trajectories in this system are therefore geodesic lines, which could be practically determined by stretching a string between the two given points so that its length is a minimum. The method is equivalent to a projection of the meridian plane in such a way that each element of length is magnified by an amount proportional to the local value of the refractive index. If the shape of elements of area is to be preserved—that is, if the projection is to be isomorphic—then the refractive index must be independent of direction at any given point; such isotropic conditions exist in the electrostatic field, and can be introduced into the magnetic lens by transforming it into the equivalent electrical lens (section 35). The appearance of the potential distribution in a two-cylinder lens in the new system is shown in Fig. 78. As the potential and therefore also the refractive index decreases with radius in any cross-section, higher values of dS must exist at the axis than near the electrodes, and this is only possible if the S -surface is curved in the way indicated. The equipotential and force lines then take the form shown by the families of curves drawn in the surface.

The Ideal Lens. An ideal lens is a system such that all trajectories emanating from a given point intersect again in a single point. In the transformed surface this can only be so if it has the form of a sphere, in which case the image of any point (O) in the surface will be the farther end (I) of the diameter through that point. Object and image will be at opposite poles of the sphere, and the trajectories will all be great circles

across its surface. The physical distribution of refractive index (or space charge) that would correspond to such an ideal lens of uniform refractive index can be immediately found by reprojecting the spherical surface

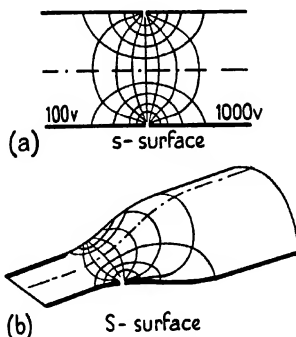


FIG. 78. Transformation of potential distribution in two-cylinder lens. (a) Meridian plane of lens; (b) corresponding *S*-surface. (Gabor.)

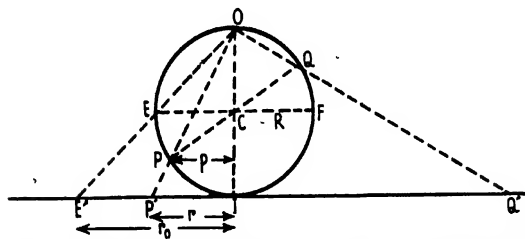


FIG. 79. Projection of ideal potential distribution (*S*-surface) on to *s*-plane. (Gabor.)

on to a plane. Any isomorphic projection will suffice, and the stereographic method is shown in Fig. 79: lines are drawn from the point *O* through the sphere to intersect a plane tangential to it at the other end *I* of the diameter through *O*. A representative line of the projection (*OPP'*), and a particular line *OEE'* through the equator are shown. The projection will result in a series of concentric circles in the *s*-plane, having centre at *I* and radius *r* given by

$$r = p \left(\frac{OP'}{OI} \right)^2 = p \frac{(2R)^2 + r^2}{(2R)^2},$$

where *p* is the normal distance of the point *P* from the diameter *OI* and *R* is the radius of the sphere. The radius of the circle through *E'*, the projection of the equator, will be $r_0 = 2R$. The linear magnification in the projection is then

$$\frac{r}{p} = M = 1 + \left(\frac{r}{r_0} \right)^2,$$

and the distribution of refractive index in the s -plane must be given by

$$N = N_0/M = \frac{N_0}{1 + (r/r_0)^2},$$

where $N(\propto \sqrt{V})$ gives the electrostatic potential at the point. The resulting distribution of refractive index with radius is shown graphically in Fig. 80a. The ideal lens (in two dimensions) proves to be a nebula of space charge of

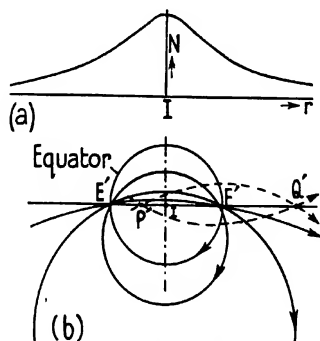


FIG. 80. Ideal lens: (a) radial distribution of refractive index (N), (b) trajectories. (Gabor.)

positive sign inside the 'equatorial' circle of radius r_0 , and negative outside it. It is impossible to achieve such a distribution of refractive index with any physically realizable combination of a magnetic field with an electrostatic field produced in the normal way by applying potentials to fixed electrodes, that is, without employing space charge.

A number of possible trajectories in the ideal lens are shown in Fig. 80b. The equatorial circle is itself a trajectory, and the other great circles of the S -system project as circles in the s -plane, as indicated. Any point P' within this circle will be imaged in a conjugate point Q' outside it, derived from the antipodal point Q of the original point P in the S -surface. The position of Q' is given by the relation $r_p r_q = r_0^2$, where r_p and r_q are the distances of P' and Q' respectively from the origin I in the s -surface.

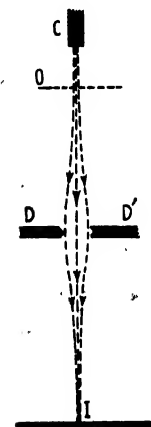


FIG. 81. Space-charge focusing. (Borries and Ruska.)

A practicable lens, of course, will not be confined to a plane in the way so far assumed. By rotating the map of the potential distribution in the s -plane about any diameter through I , we obtain a spherical nebula of space charge, which would likewise form a perfect imaging system. However, the prospect is remote of so controlling the establishment of space charge that the necessary decrease in density with

radius could be realized, apart from the difficulty of forming a point source of electrons or ions at the centre of the sphere. A slightly greater chance of success obtains for the cylindrical lens (in the glass optical sense) formed by translating the potential distribution in the s -plane parallel to itself along the normal through the centre of symmetry I . Such a space charge might conceivably be established by an inversion of the normal construction of the diode thermionic valve, the axial electrode becoming a source of positive ions and the cylindrical 'anode' being negatively charged. An alternative possibility is to use a low-pressure electrical discharge for the purpose, and indeed Borries and Ruska investigated such a space-charge lens some ten years before Gabor's demonstration of its possible ideal properties. They used the arrangement shown in Fig. 81, where C is a source of electrons which illuminate an object O and are subsequently focused by the auxiliary discharge passing between electrodes D and D' . The images they obtained of a wire graticule were recognizable but much distorted.

FURTHER READING

1. General Theory of Aberrations:

- Scherzer, *Z. Phys.* **101**, 593, 1936.
 Gray, *Bell Syst. Techn. J.* **18**, 1, 1939.
 Ramberg, *J. Opt. Soc. Amer.* **29**, 79, 1939.
 Voit, *Z. f. Instk.* **59**, 71, 1939.

2. Experimental Investigation:

- Klemperer and Wright, *Proc. Phys. Soc.* **51**, 296, 1939.
 Gabor, *Television*, **11**, 206, 1938.
 Diels and Wendt, *Z. techn. Phys.* **18**, 65, 1937.

3. Spherical and Chromatic Aberration:

Electrostatic Lenses:

- Rebsch, *Ann. Phys.* **31**, 551, 1938.
 Plass, *J. Appl. Phys.* **13**, 49, 1942.
 Ramberg, *ibid.* 582.

Magnetic Lenses:

- Glaser, *Z. Phys.* **117**, 285, 1941.
 Dosse, *ibid.* 722.

4. Aberrations of the Image Tube:

- Henneberg and Recknagel, *Z. techn. Phys.* **16**, 230, 1935.
 Morton and Ramberg, *Physics*, **7**, 451, 1936.

5. Space Charge Focusing:

- Gabor, *Electronic Engng.* **15**, 295, 328, and 372, 1942-3.
 Borries and Ruska, *Z. Phys.* **76**, 649, 1932.

CHAPTER VI

THE PRODUCTION OF ELECTRON BEAMS

45. *Electrons in Metals*

IN previous chapters the fundamental behaviour of electrical and magnetic lenses has been discussed in some detail; the action of an isolated lens on an electron beam should now be clear. Succeeding sections will be devoted to a brief outline of some of the more important practical applications of these principles, but before passing into the field of applied electron optics we shall indicate the means by which the essential electron beam is obtained. Starting from the standpoint of the modern theory of metals, the various types of electron emission are discussed and in each case a practical device utilizing it described. Fuller treatment will be found in more specialized books, some of which are listed at the end of the chapter.

The picture provided by the new quantum theory of the structure of metals is that of an ordered arrangement of atoms, each of which contributes one (or sometimes more) electrons to the so-called 'electron-gas' that circulates comparatively freely among the atoms. The remaining electrons are bound to the nucleus, about which they move in orbits or shells of differing radii; it is the outermost, or valence, electrons which are released and become in a sense common to all the atoms of the metal lattice. When an external electric field is applied between two points in the metal, these free electrons (being negatively charged) move in the direction of increasing potential and constitute a current. Since, on the average, each atom gives up one electron to the common pool, the number of electrons per cubic centimetre (the 'electron concentration') is of the order of 10^{22} . Their drift velocity in the direction of the applied field need, therefore, be only of the order of a few cm./sec. in order to constitute a current of practical significance (\sim an ampere). On the other hand, the velocity of random motion of the free electrons among the atoms is considerably greater than this, being much higher than the thermal velocity of gas molecules. The total charge thus carried, in one direction, across any boundary surface within the metal is found to correspond to a current of the order of 10^{13} amp./sq. cm. The random current in the opposite direction practically compensates it, although the varying difference or statistical fluctuation cannot be neglected when the conductor is carrying a very small directed current which is later to be subjected to high amplification; it then forms a background 'noise' which sets a limit to the ~~smallest~~ signal that can be significantly amplified.

Although these electrons are 'free' in the sense of not being attached to a particular atom, they are under the influence of the attractive field of the lattice as a whole and hence are not free to escape from it into the gas or vacuum above its surface. For this to occur an electron must acquire

additional energy, which may be derived from the lattice itself, when the electron may be said to be expelled, or from an externally applied field which extracts it from the solid metal; in special circumstances the energy may be supplied by incident radiation or corpuscles. Quantitative discussion, therefore, proceeds in terms of the energy relations of the lattice and its free electrons.

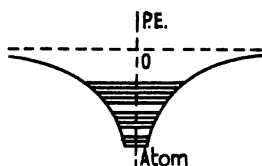


FIG. 82. Potential energy of an electron near an isolated nucleus. (Horizontal lines = energy levels.)

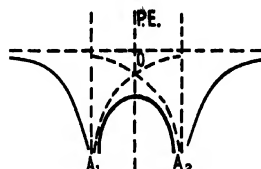


FIG. 83. Potential energy of an electron along a line through two adjacent atoms (A_1 and A_2).

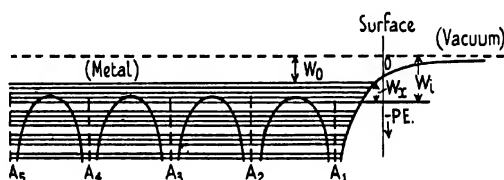


FIG. 84. Potential energy of an electron along a line through a row of atoms in a metal, showing energy bands; the Fermi-Dirac distribution applies to the topmost band only.

Considering first an isolated atom, we imagine a valence electron to be extracted from it and conveyed to infinity. Not being in a lattice, the atom will not automatically release the electron, which will be under the influence of the ordinary Coulomb force of electrostatic attraction; therefore work will have to be done in order to separate it from the rest of the atom. If the nucleus only were in question, the attractive force would be inversely proportional to the square of the distance of the electron from the nucleus at any time. On account of the other electrons around the nucleus these relations are in fact more complicated, but the potential energy (P.E.) of the electron with respect to the atom can be approximately expressed by the inverse square law (Fig. 82); the horizontal broken line indicates its potential at infinite distance. If now two neighbouring atoms (A_1 , A_2) are considered, the potential energy curve has the form shown in Fig. 83; the presence of the second atom decreases the potential on that side of the first atom, making it slightly easier to extract the electron from the latter. When there is a continuous lattice of atoms, this mutual effect is even more marked and results in a series of potential humps between the atoms, at a still lower level below the reference line (Fig. 84). At the end of the row of atoms the potential rises in much the same way as near an

isolated atom, indicating that an electron is here under a considerable attractive force from the lattice as a whole. On this picture it is clearly difficult to give a precise meaning to the 'surface' of the metal; for convenience it may be set at a distance from the end atom equal to half the separation between neighbours in the lattice (vertical broken line).

Within each atom the electrons do not have a continuous energy distribution. According to the quantum theory only a strictly limited number of states is available, and these occur in bands within which are several permitted levels and outside of which there are 'forbidden' zones. In the lower levels, corresponding to a closer proximity of the electron to the nucleus, these bands are narrow and widely spaced; the higher bands are much broader and contain more levels, and may overlap each other. In Fig. 82 this level system is indicated by full horizontal lines, confined within the potential curves in order to show that the electrons are bound to the isolated atom. Each level may be occupied by no more than two electrons, of opposite spin, as required by Pauli's exclusion principle.

In the complete lattice the same limiting conditions prevail. There are a restricted number of states available, grouped in bands, but these bands now extend throughout the whole lattice and each contains as many energy levels as there are atoms in the lattice (full lines, Fig. 84). For, according to quantum mechanics, an electron may jump from a given state in one atom to a state of equal energy in another; however, the probability of such a jump ('tunnel effect') decreases exponentially with the height and breadth of the potential hump which has to be traversed between the atoms.

In normal circumstances there are more than enough electrons present to fill the lower levels, and some occupy levels in the band which rises above the top of the potential humps; indeed, some may have energies of several electron volts higher than this. It is such a distribution of states that determines the special properties of metals; these higher levels provide the electrons which circulate freely through the lattice. The maximum energy of these electrons, and the probability of escape from the metal, is found not by consideration of individual electrons and their levels but by treating them as a 'gas' and applying statistical methods as in dealing with an actual gas. In normal kinetic theory, however, a Maxwellian distribution of energies amongst the particles is assumed, where the majority have values around the average (\bar{W}_a) (Fig. 85). In the electron gas, as mentioned above, all the lower states are normally filled, so that the energy distribution must be very different. The Fermi-Dirac treatment gives its form: the *probability* ($P(W)$) that a given state will be filled is unity for all states up to a definite maximum energy, when the lattice is at absolute zero (0°K.). The probability curve therefore has the rectangular form shown in Fig. 86a; the average number ($N(W)$) of electrons having energies lying within a given energy range (dW) will be different, however, as there are many states with energies in this range. $N(W)$ proves to be proportional to the root of the energy, and hence the

distribution of energies amongst the electrons has a parabolic form (Fig. 86b). Both curves fall sharply to zero at the maximum value of the energy W_x , when the metal is at 0°K. and this level is naturally below the potential of the electron at infinity, W_i , which is termed the inner potential of the lattice and is of the order of 10 V. (Fig. 84). Except by the tunnel effect, electrons can only escape from the metal if they acquire a total

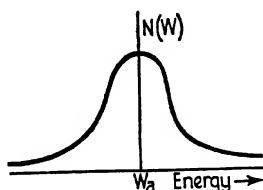


FIG. 85. Maxwellian distribution of energies amongst particles.

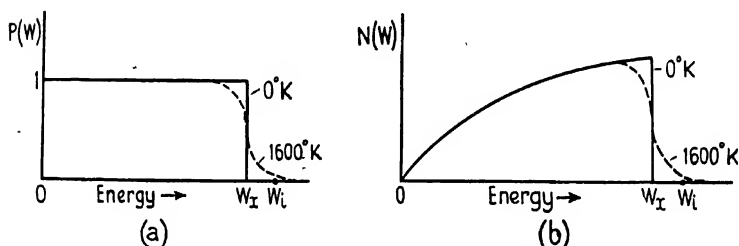


FIG. 86. Fermi-Dirac distribution for electrons: (a) Probability for a given state. (b) Frequency of energy distribution.

energy greater than W_i . The work W_0 that must be done to extract an electron in this way is therefore given by

$$W_0 = W_i - W_x = V_0 e, \quad (\text{VI.1})$$

when expressed in terms of the voltage V_0 necessary for the purpose; V_0 is termed the work function (at 0°K.) of the surface.

The energy needed to surmount the potential barrier is comparatively large, V_0 being usually several electron volts. It may be supplied, or its amount considerably reduced, in any of the following ways:

(a) *Thermal Excitation.* If the temperature of the metal is raised above 0°K. there is at first a negligible effect upon the energies of the free electrons: they are in energy levels high above the thermal energy of the lattice atoms, and therefore do not exchange energy with them. At room temperature the maximum energy W_x is about 6 electron volts, being still several electron volts below the value of W_i . At high temperatures, however, the thermal energy of the lattice approaches that of random motion of the electron gas, and the two tend to an equilibrium. There will now be a roughly Maxwellian distribution of energies amongst the free electrons in

the higher states, and the curves of Figs. 86*a* and *b* will be rounded off about the limit W_x , as shown by the broken lines (for $1,600^\circ \text{ K.}$). Some electrons will now have an energy above the inner potential and will be ejected from the surface. This *thermionic emission* is appreciable from a pure metal only at temperatures of the order of $2,500^\circ \text{ K.}$, but oxidation of the surface, or contamination with another metal, lowers the working temperature very considerably (see next section).

(*b*) *Lowering of the Barrier.* Instead of supplying the electron with energy enough to raise it to the level (W_i) of the potential barrier at the surface, the latter may be lowered in order to allow the more energetic electrons to escape. To do this it is in principle only necessary to apply a counter-field at the surface sufficient to reduce W_i towards the value W_x and thus reduce the work function. The field required to extract electrons in this way proves, however, to be very great owing to the fact that the barrier extends over approximately the radius of an atom only. The field strength at the surface is therefore of the order of a few volts per Angstrom unit, and the counter-field would have to be similar in magnitude: complete lowering of the barrier would demand a field of millions of volts per cm. On account of the tunnel effect, however, it is sufficient to establish a positive field of only some thousands of volts per cm. in order to get an appreciable emission. Further discussion of such 'field' or 'cold' emission takes place in section 48.

(*c*) *Absorption of Radiant Energy.* If radiation falls on the surface of the metal, energy will be transferred to the free electrons in amount depending on the wave-length (λ) of the radiation and the reflection coefficient for this wave-length. According to the quantum theory, absorption takes place in discrete amounts of energy W given by

$$W = h\nu, \quad (\text{VI.2})$$

where h is Planck's constant and ν is the frequency of the incident waves. When the latter corresponds to a wave-length of less than $10,000 \text{ \AA}$, the energy quantum is of the order of electron volts and therefore approximates to the work function V_0 . It follows that infra-red and visible light, as well as X-rays and gamma-rays, will be capable of exciting electrons to levels high enough for them to escape from the metal. Such *photo-emission* is of very low efficiency, however, since only a portion of the radiation is absorbed and a much smaller proportion of the excited electrons succeeds in reaching the surface and escaping. It is further considered in section 50.

(*d*) *Absorption of Fast Particles.* Energy will also be transferred to the lattice if the incident beam is composed of corpuscles, and not of radiation. When the corpuscles are electrons, the transfer of energy to the free electrons is more efficient than in the case of larger particles, such as positive ions. A proportion of the lattice electrons will be raised to energies sufficient for escape from the lattice, and also some of the incident beam will be scattered back from the surface. The energy of each absorbed particle may be shared among several lattice electrons and, when the

incident beam is of high velocity, several electrons may be emitted for each of those incident. Certain specially treated surfaces show a return current more than ten times greater than the primary current. *Secondary emission*, as it is called, is consequently of value in amplifying small currents (see section 52).

Each of these four general methods of obtaining electron emission finds application in practical devices, and occasionally they are used in combination. Thermionic emission is most widely used for obtaining electron beams in radio valves, cathode-ray tubes, electron microscopes, and similar apparatus. In some special circumstances, however, it yields in advantage to secondary emission and even to cold emission; for other purposes it is entirely inapplicable, as, for instance, in the conversion of an optical into an electron-optical image, and one of the other methods must be employed (in this case, the photo-electric effect).

46. *Thermionic Emission*

A temperature of the order of $2,500^{\circ}\text{K.}$ is needed, as stated above, before electrons in any number acquire sufficient energy to escape from the surface of a metal. Very few metals remain solid at this temperature, and it is usual to use tungsten or molybdenum plates or wires in the investigation of thermionic emission from pure elements. In order to prevent the accumulation of emitted electrons as a space charge around such a cathode it is necessary to insert a positively charged electrode, or anode, in its neighbourhood; for present purposes we may take this to be a cylindrical anode surrounding a straight tungsten wire on the axis. The increase of current with temperature is found to be exponential in form (Fig. 87). By assuming that the emission from the surface could be treated as an evaporation of particles having a Maxwellian energy distribution, Richardson found that the maximum current per square cm. (i_s) obtainable at a given temperature T depended also on $T^{\frac{1}{2}}$. In fact, the Fermi-Dirac distribution applies, and, moreover, the number of free electrons per molecule is not constant, as Richardson assumed, but increases with temperature. Therefore the current must depend on a higher power of T , but direct experimental verification is made difficult by the preponderating effect of an exponential term on the value of the current. From quantum mechanical considerations, however, Nordheim has obtained the modified form of Richardson's equation:

$$i_s = AT^{\frac{3}{2}}e^{-V_0/kT}, \quad (\text{VI.3})$$

where k is Boltzmann's constant (1.371×10^{-16} ergs per degree), V_0 is the work function already defined (in ergs per electron), and A is another constant characteristic of a given surface. By accurate measurement of the variation in emission with temperature it is then possible to determine the value of the work function. For tungsten, after careful outgassing, it

has the value 4.52 V. and the order of magnitude of the emission (i_s) is 10 ma./cm.² at 2,200° K., rising to 100 ma./cm.² at 2,400° K.

It is inconvenient, and wasteful of power, to use such high temperatures in the cathode. Fortunately it is found that contaminating the tungsten surface with another metal lowers its work function below that of either

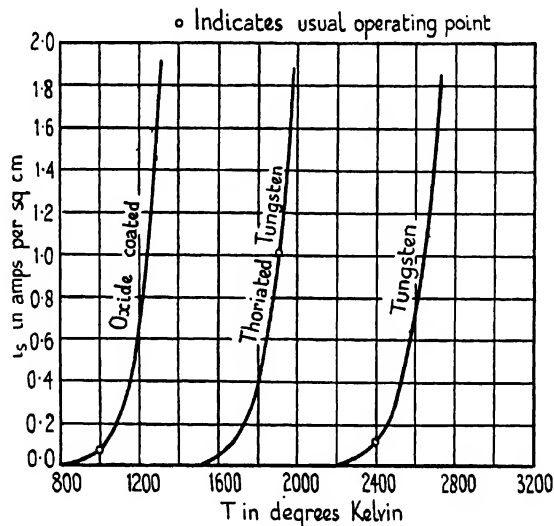


FIG. 87. Variation of thermionic emission with temperature.

metal separately. Thus a thoriated tungsten wire has a work function of only 2.6 V., compared with 4.52 V. for tungsten and 3.35 V. for thorium. Other electropositive metals have the same effect, apparently setting up a counter-field that lowers the potential barrier, whereas electronegative atoms have the reverse effect and raise the work function. The efficiency of emission is considerably increased: it is 3 ma./watt for tungsten at 2,400° K., and as much as 100 ma./watt for thoriated tungsten at 1,900° K.

The working temperature may be much further reduced if oxygen is admitted to the surface at some stage in the 'activation'. The precise nature of the processes involved is still in doubt, but they are found to lower the work function to the order of 1 V. and result in very ready emission of electrons. Caesium deposited on oxidized tungsten then yields 350 ma./cm.² of surface at 1,000° K. only, and surfaces of calcium or barium oxide, derived from the respective carbonates by decomposition, are equally powerful. At this temperature, however, the efficiency is not proportionately so great, being about 10 times that of a pure tungsten surface. Oxide-coated cathodes of this 'dull-emitter' type are now in almost universal use in electron-optical apparatus; the chief exception is the electron microscope, in which tungsten filaments are required in order to give a point cathode. The variation in emission with temperature

for oxide-coated cathodes is shown in Fig. 87, and has the same form as that for pure metals. Their properties depend rather critically on the activation procedure, details of which will be found in the technical treatises.

47. The Electron Gun

Thermionic emission from an oxide-coated cathode provides a source of electrons delivering currents of the order of an ampere per cm.² In the cathode-ray tube and many other devices the technical problem is then

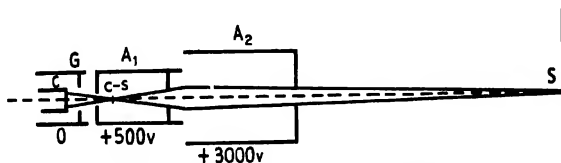


FIG. 88. Electron gun with second anode, showing beam cross-over ($c-s$).

to direct the emission into a beam which shall finally form a minute spot on a viewing screen. The arrangement used for this purpose is called an electron gun, the first essential of which must be an immersion lens of the type discussed in section 27. It is obvious from Fig. 43, however, that, as this system sharply converges the emitted electrons into a small spot (or cross-over) at some point within the anode A , the beam must on this account issue with a strong divergence. In order that a final image shall be formed, a second focusing system must be introduced beyond A . In early types of cathode-ray tube a magnetic coil was used, or reliance was simply placed on the focusing action of the ions produced in the residual gas in the rather low vacuum prevailing. In the 1930's much use was made of two- or multi-cylinder electrostatic lenses, but latterly the trend has been back to the simpler operation possible with a magnetic focusing coil.

(a) *Electrostatic guns.* The addition of a second anode A_2 forms a two-cylinder lens with the gun anode A_1 , and if the potential ratio $V_2:V_1$ is made much greater than unity there will be a strong converging action on the beam. This arrangement was first proposed by Zworykin and is shown in Fig. 88. In order to reduce the size of the final image, diaphragms are now fitted to the grid and first anode, as discussed in section 27 (p. 69). The second anode is usually of greater diameter than the first, and may also be fitted with limiting diaphragms the function of which is now to trap stray electrons. In later developments more than one extra anode has been used, the second system being a symmetrical or twin-aperture lens, or even more complex. In all cases the focusing action is basically the same.

The emitting cathode C has an area of many square millimetres, in order to achieve a high beam current, and it is preferable therefore to use as

object for the second lens the much smaller cross-over CS in front of the cathode. This acts as a virtual cathode, and corresponds optically to the pupil of the system. It was shown by Ruska that in a simplified gun the demagnification in the cross-over, with respect to the cathode, is given by the expression

$$\frac{y}{y_0} = \frac{2}{\sin 2\theta} \sqrt{\frac{V_c}{V_1}}, \quad (\text{VI.4})$$

where y and y_0 are the radii of cross-over and cathode respectively, V_c and V_1 are the voltages of thermionic emission and the first anode, and θ is the semi-vertical angle of the cone of rays forming the beam at the cross-over. In practical circumstances this demagnification will be of the order of 1:10, so that the cross-over will be only a fraction of a millimetre in radius. The potential ratio between the two (or more) anodes is then so arranged that an image of this cross-over is projected on the fluorescent screen (S) at the end of the apparatus; for a given tube the object-image distance is fixed and hence a fixed voltage ratio is also demanded. Owing to the interposition of a deflecting system and the need for a subsequent path long enough to give a large sweep on the screen, the image distance must always be longer than the object distance. The final image of the cross-over is therefore enlarged rather than diminished, and it is for this reason that particular attention is given to making the cross-over as small as possible in the first place.

The cross-over, however, cannot be a point because of the mutual repulsion of the electrons and of the aberrations of the immersion lens system. The former factor becomes important when attempts are made to increase the current density in the beam, as is essential for a bright picture in reproducing a television transmission. Both the spherical and chromatic errors of the lens are appreciable, the latter arising from the variation in V_c , the energy emission of the electrons. But even were the lens perfect and V_c constant, the current density in the cross-over would be limited, as is evident from the fact that V_c enters into the expression (VI.4) for the size of the latter. On the assumption that the emitted electrons have a Maxwellian distribution of energies, the problem has been investigated by D. B. Langmuir, who finds that the maximum current density obtainable in the cross-over (i_x) is given by

$$i_x = i_0(eV_1/kT + 1)\sin^2\theta, \quad (\text{VI.5})$$

where k , V_1 , and θ have the meanings already assigned, T is the absolute temperature of the cathode surface, and i_0 the current density (in amp./cm.²) at the cathode. A similar expression holds for the current density in the final image spot, with appropriate insertion of the final voltage and angular size of the beam. Pierce has confirmed these results by a more fundamental discussion, using the methods of statistical mechanics instead of the optical argument followed by Langmuir.

Langmuir's expression shows that the current density in the cross-over may be increased by increasing the potential of the region in which it

falls (V_1), by decreasing the working temperature of the cathode (and hence the energy of emission, V_e), by increasing the aperture of the cone of rays focused or, of course, by increasing the output of the cathode itself. Assuming the cathode temperature to be $1,160^\circ \text{K.}$ and its emission to be 1.0 amp./cm.^2 , the calculated maximum value i_x in the cross-over is 10 amp./cm.^2 for $\theta = 0.032$ and 100 amp./cm.^2 for $\theta = 0.10$, when the anode voltage $V_1 = 1,000 \text{ volts}$; for an accelerating potential of 10 kV. ,

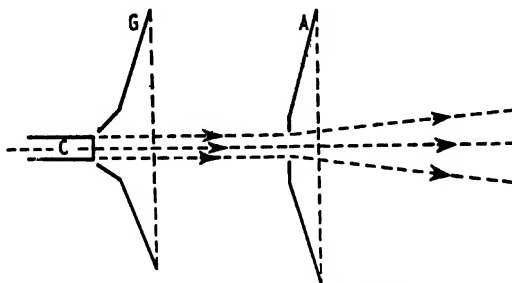


FIG. 89. Electrode form producing parallel beam. (Pierce.)

the calculated current densities will be 10 times greater in each case. The value in a given gun is always considerably less than calculated, owing to aberrations of the lens and mutual repulsion of the electrons. More intense beams can only be produced by improving the cathode, so that the output is greater and the temperature lower, or by attention to the lens, which must carry a higher voltage and be well corrected so that beams of large aperture may be adequately handled. The second lens system must be similarly perfected so that all the electrons from the cross-over are focused into the final spot. At present intensities on the screen of over 10 amp./cm.^2 have been reached, but there is no doubt that considerable further improvement is possible.

A direction in which progress is now being made is in designing the form of lens apertures so that the electron paths are truly rectilinear. Pierce has carried out the necessary calculation of the equipotential surfaces for beams of differing symmetry, on the assumption that they are under no other influence. He thus neglects the effect of space charge as well as that of the electromagnetic mutual attraction of the electrons, both of which will tend to make the beam more compact; he neglects also the velocity of emission of electrons from the cathode. The field distribution having been found for (say) a beam of cylindrical form, any two equipotential surfaces may be chosen for replacement by electrodes of the same shape and bearing the same relative potentials. In this instance the resulting forms are saucer-shaped, as shown in Fig. 89; the beam diverges slightly beyond the anode, which acts as a weak aperture lens. Beams of other (constant) cross-section and of wedge shape have been investigated in the same way, as also the special case of electrons directed towards

the centre of a sphere, and appropriate electrode forms indicated. This method of approach promises a better approximation to perfect focusing than the empirical methods usually adopted in arriving at the shape of apertures for electron lenses. Pierce claims to have obtained a current density (i_x) in the cross-over greater than one-half the maximum value given by theory (VI.5), with currents of the order of 100 ma.

(b) *Magnetic Focusing.* The use of a magnetic coil for concentrating the divergent beam provided by an immersion system was adopted by Farnsworth in 1934. It was really a reversion to the method used in early

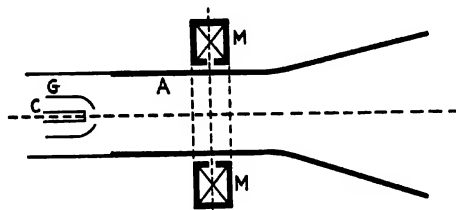


FIG. 90. Electron gun with magnetic focusing coil and wall-anode.

cathode-ray tubes employing a primitive cathode and simple anode as gun, and a magnetic focusing coil. His tube had also a grid, and a conducting coating on the inner surface which acted as a second anode. A few years later the first anode was eliminated, only the grid and the graphitic or metallic coating being retained. The system was thus reduced to the simplest possible form (Fig. 90), and yet functioned as well as any other model. In addition, the external mounting of the magnetic coil M allowed of adjustment at any time after assembling and sealing off the tube, an advantage which is absent from the electrostatic type.

A cross-over is formed in the usual way by the immersion system of cathode, grid, and anode, and is then imaged on the viewing screen by the magnetic field. The current density in the cross-over is given as before by expression VI.5, and the density at the final image by a similar relation involving the magnetic field in place of the accelerating potential. The chief disadvantage is the spherical aberration of the magnetic lens. This, however, is to some extent offset (as in the electrostatic type) by the natural depth of focus associated with all such beams of narrow aperture. The production of a point image is opposed by the mutual electrostatic repulsion of the electrons, but this in turn is partly counteracted by the mutual attraction arising from their magnetic fields, just as any parallel conductors carrying current in the same sense experience an attraction. As a result the beam tends towards a uniform cross-section, in which the electron paths are nearly parallel, as it approaches the image point. This great depth of focus renders spherical aberration less harmful and at the same time makes the position of the imaging screen less critical. Nevertheless, the problem of spherical aberration remains serious for beams of large

aperture, and the achievement of higher intensities depends again mainly on the correction of the lens for this defect.

Tubes incorporating both types of secondary focusing are also in use, and one of the most successful is that due to Law. It embodies four anodes in the immersion system, three of them simple apertures at successively higher potentials (Fig. 91), so that the cross-over is finally formed in a

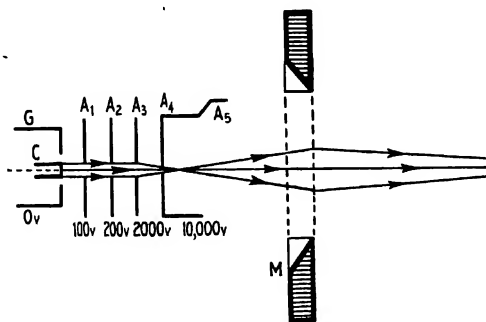


FIG. 91. Law's electron gun.

region of very high potential (10,000 V.). The size of the cross-over is thus reduced and the current density in it considerably increased (cf. VI.4 and VI.5). The magnetic coil (*M*) then projects on to the viewing screen an image of the cross-over as a small spot of high intensity; it is specially designed to have low spherical aberration, this being claimed to be negligible over an aperture of 6 mm. The tube is accordingly employed in television for producing very bright images for subsequent optical projection on to a larger screen. It will be noticed that the grid is at the same potential as the cathode, and the final anode is connected to a conducting coating (*A₅*) on the wall of the tube, which ensures that the whole of the image space is at the same high potential.

(c) *Gas Focusing.* The presence of residual gas in a cathode-ray tube does not, as might have been expected, scatter an electron beam into a diffuse bundle, but is found to exert a pronounced concentrating effect upon it. Indeed the early forms of tube depended to a large extent on this phenomenon for the degree of focusing obtained. In the Braun tube (p. 3), for instance, the cold emission from a simple anode is to some extent directed by the strong field, but would give a spot of negligible intensity on the screen were it not for the effect of the residual gas. In later tubes a grid as well as an anode was employed, so that a simple immersion system existed and a rather ill-defined cross-over was formed within it. It was found in practice that the final spot was smaller in size if the tube was not highly evacuated, but only to a pressure of the order of 10^{-3} or 10^{-4} mm. Hg, for which purpose an inert gas such as helium or argon was introduced into it after preliminary evacuation. The gas, acting as a

'long lens', maintains an almost parallel beam from the anode to the viewing screen.

The details of the process are still not clear, but it depends essentially on the formation of positive ions in the gas; the energy of the electrons being well above its ionizing potential, ions are readily formed by collision. These positive ions drift towards the cathode under the influence of the applied field, but being much heavier than electrons they move much more slowly. Hence there are always a considerable number within the electron beam, forming a loose 'core' which binds it together by electrostatic attraction. A state of dynamic equilibrium is set up, which in special circumstances may result in a parallel beam of as much as a metre in length.

Under such pressure conditions the beam partakes of the nature of a discharge, and the main source of electrons is the bombardment of the cathode by positive ions. Indeed, in the cold cathode tube this is almost exclusively the case; if the gas pressure is reduced below 10^{-3} mm. Hg, the discharge ceases and the beam current becomes negligible. But so long as the discharge continues, the ionization of the gas results in a secondary supply of electrons, which may be considerable if the accelerating potential is high enough (as in most cold cathode tubes) for multiple ionization to occur. If the potential rises too far, or alternatively, if the gas pressure rises, the discharge passes over into an arc carrying a large current. On the other hand, at potentials below 300 volts the amount of ionization is too small to form a coherent beam, and no sharp spot is obtainable.

The gas-focused oscillograph has now passed out of use for a number of reasons, chief among which are the low deflexion sensitivity and the early defocusing of the beam with increasing frequency of deflexion. The cathode is also rapidly disintegrated by the heavy positive-ion bombardment, and this process may be further accelerated by unskilled handling of the tube. Nevertheless, gas-filled tubes find a number of uses in other directions.

We may, therefore, sum up the discussion of electron guns by saying that the primary essential for producing a beam is an immersion electrostatic lens, followed by a second lens which focuses a sharp image of the cross-over on to the viewing screen. The second lens is often one or other of the electrostatic types, but increasingly of recent years a magnetic focusing coil has been preferred; the early method of gas-focusing is now rarely used. The same general conclusion applies to the electron microscope and other special devices as well as to the cathode-ray tube.

48. *Field Emission*

As seen above, thermionic emission takes place only at high temperatures, and is normally negligible for even an oxide-coated cathode at less than 800° K.; the free electrons in the surface have insufficient energy to escape over the potential barrier. The application of a high positive field

in its neighbourhood, however, can produce considerable emission even at room temperature and under special conditions this may be much greater than the maximum thermionic emission. The processes at work are quite different from those operating in the 'cold-cathode' tubes mentioned above, where the cathode yields electrons under positive-ion

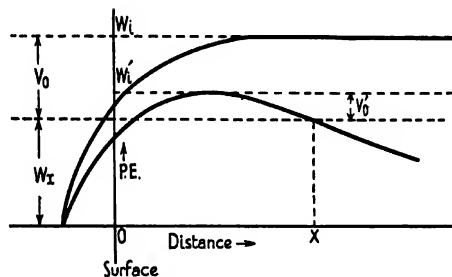


FIG. 92. Lowering of potential barrier by an external field.

bombardment; true field emission occurs in conditions of the highest vacuum.

The essential action of the field, as indicated in section 45, is to lower the potential barrier at the surface so that the effective work function V_0 is reduced. Since the latter enters exponentially into Richardson's equation for the emission current density (VI.3), a small reduction in it results in a considerably increased emission if the value of T is not too small. At the same time, the lowering of the barrier allows the tunnel effect to come into play, and this can give rise to a large emission whatever the temperature may be.

Schottky Effect. The effect of the applied field, of strength E , may be supposed to increase directly with the distance outwards from the surface. The potential barrier, normally of height W_i , will accordingly be modified as shown in Fig. 92, having now a maximum height W_i' . The difference represents the decrease in the work function for an emitted electron:

$$W_i' - W_i = \Delta V_0,$$

its new value V_0' being given by

$$W_i' - W_x = V_0'. \quad (\text{VI.6})$$

The insertion of this value in Richardson's equation will therefore give a higher result for the current density. It may be shown that, if i_s is the density calculated in the absence of the positive field, then the new value i_s' is given by

$$i_s'/i_s = e^{4.403 \sqrt{E/T}}, \quad (\text{VI.7})$$

an equation first derived by Schottky, after whom this type of emission is often called the 'Schottky effect'. Since i_s' is also governed by Richardson's equation (VI.3), the increase in emission will only be appreciable for

large values of T . In calculating its amount the value of E at the surface must be employed, as determined by the form of the electrodes. In the ordinary diode, for instance, we have a cylindrical cathode wire surrounded by a concentric anode; the dependence of the field on radial distance is therefore given by the logarithmic relation of section 12(c), and will be much larger near the wire than if plane parallel electrodes were involved. If in this case the filament is a tungsten wire of diameter 0.1 mm. inside an anode of 1 cm. diameter maintained at a potential of 44 V., the field intensity close to the wire will be 1,900 V./cm. and not 88 V./cm. At a temperature of 2,000° K. the increase in emission calculated from VI.7 would be 10 per cent. only. At an anode potential of 4.4 kV., however, the increase would be 8,000-fold over the simple thermionic emission at this temperature. In the electron microscope, for instance, it makes an important contribution to the beam current.

Tunnel Effect. The tunnel effect is very different in nature, and depends on the finite probability (in wave-mechanical theory) that an electron on one side of a potential barrier may pass across it if there is an unoccupied state of the same energy on the other side. In the absence of a positive field at the surface of the cathode the potential barrier has no 'other side', but continues indefinitely at the height W_i (Fig. 92). When the barrier is modified into the extended hump depicted, it will at some distance from the surface fall below the energy level of the topmost band in the metal which is occupied by free electrons, and a finite probability will exist that they will jump the barrier. The higher the applied field the closer to the surface is the point (X) at which the barrier drops to this critical level, and the greater the probability of 'tunnel' emission. It has been shown by Fowler and Nordheim that this *auto-electronic* current is given by an analogous equation to Richardson's,

$$i_s = A_s E^2 e^{-b/E}, \quad (\text{VI.8})$$

where E is the field strength at the surface and A_s and b_s are approximately constant, the latter depending on the work function V_0 . The equation gives good agreement with experiment for carefully cleaned pure metals, but, like Schottky's equation, is not obeyed by contaminated surfaces. *Auto-electronic* emission is thus independent of the temperature and increases rapidly with field strength, being determined by the rate of arrival of free electrons at the surface, which, as mentioned earlier, is equivalent to a uni-directional current of the order of 10^{13} amps./cm.² Hence, although the proportion allowed through the barrier by the tunnel effect is very small, the emission may already exceed the normal thermionic emission at quite moderate field strengths and at room temperature. Contamination of the surface with another metal lowers the work function, and therefore b_s , and increases the field emission further. By applying intermittent fields of great strength, intensities of the order of 1,000 amps./cm.² were obtained by Slack and Ehrke for the purpose of generating X-rays of high intensity.

49. Field Emission Microscope

A valuable and extremely simple application of the principle of field emission has been to the development of a diode system in which a greatly enlarged image of an emitting-point cathode is projected on to a fluorescent screen on the anode. Although it has no optical system of focusing, it is effectively a microscope and has been named the 'point-projector microscope'. As originally constructed by Johnson and Shockley it was cylindrical in form, with an emitting filament on the axis. Benjamin and Jenkins have used it for investigating the emission from cold point-cathodes of tungsten, molybdenum, and nickel, in the spherical arrangement of Fig. 93. The cathode (*C*), attached to a heating wire, is fixed at the centre of a highly evacuated glass bulb (*B*), coated with a fluorescent screen (*S*) over the upper hemisphere and with a broad graphite layer (*G*) over most of the lower hemisphere. A potential difference of approximately 10,000 V. is applied between the cathode and this conducting layer, the shape of which is such as to produce a field that is effectively radial near the point *C*; it is not necessary for the whole envelope to be conducting in order that this condition should be fulfilled. *C* may be a tungsten point finely etched in sodium nitrite to a radius of the order of 1 micron, being roughly hemispherical in form at the end. Under the influence of the high field strength existing near a surface of this rapid curvature, electrons are emitted along rectilinear paths and continue so through the comparatively weak outer field until collision with the screen. The emitting areas of the cathode are therefore truly imaged in a visible pattern, at a magnification given by the ratio of the radius of the point to that of the screen. The distance *CS* is about 10 cm., so that the magnification is of the order of 10^5 . Hence surface areas containing only a few hundreds of atoms will form visible images if the emission from them is bright enough. Dosse and Müller give 17 Å as the theoretical limit of the resolving power of the projection microscope in its present form; less than 30 Å has been claimed in practice.

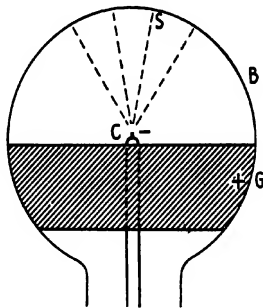


FIG. 93. Point-projector electron microscope. (Benjamin and Jenkins.)

This comparatively simple instrument is of great value for investigating the distribution of emission over the surface of cold cathodes, its modification by the addition of contaminating metals and oxides, and its variation with temperature. The main conclusion is that emission from a clean metal point occurs preferentially along the raised edges between the main crystal planes which are produced by the etching process (Plate III *a*). On heating, the emission is observed to spread over the surface in a regular manner (Plate III *b*); the lines of flow of the surface atoms, well below the melting-point, can be observed on the screen. The formation, or 'building-

up', of a cathode can thus be followed step by step; the final state is such that the emission is greatly enhanced. Experiments with adsorbed layers of barium, thorium, and other metals show that these migrate at even lower temperatures and produce a greatly increased emission, as is to be expected from the lowering of the work function. Migration of thorium on tungsten commences at 870°K. , and a fully formed point is shown in Plate III c. On further heating, the emitting layer contracts either by diffusion inwards or by evaporation of thorium (Plate III d). The complete lack of emission from certain parts of the surface is still not satisfactorily explained, however.

The projection electron microscope is a comparatively new research instrument and seems likely to find many applications in chemical as well as physical problems. It offers a means of following surface reactions and adsorption conditions, as well as of investigating both the practical and fundamental problems of cold emission. The knowledge thus obtained will in turn allow of this effect being put to other uses; the production of high intensity pulses of X-rays has already been mentioned as one of them.

50. Photo-electric Emission

When a beam of light is incident on the surface of a metal M (Fig. 94),

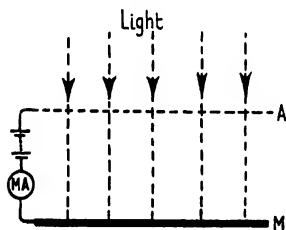


FIG. 94. Photo-electric emission.

electrons are emitted from it at low velocities. They may be collected as a current of the order of microamperes if a wire loop or gauze (A) at a small positive potential is arranged before the surface. It is found that the output of current is directly proportional to the intensity of the incident light. The variation with wave-length is complex, but for each surface there exists a maximum wave-length ('threshold') above

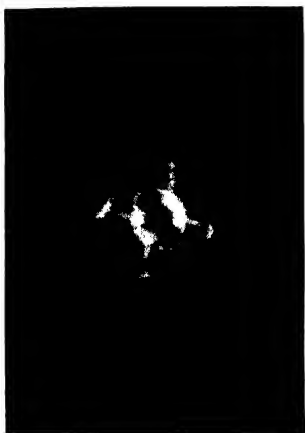
which radiation is unable to eject electrons. As in the other types of emission, contamination of the surface with another metal or an oxide causes a large increase in the current.

The experimental facts have been known for some fifty years now, and it was the first triumph of the quantum theory that it gave a satisfactory explanation where classical theory failed. The energy in the incident light is assumed to arrive in a series of discrete packets or *photons*, each bearing an amount

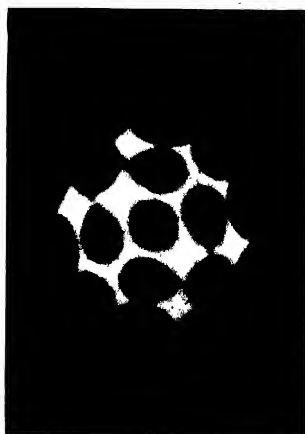
$$W = hc/\lambda = h\nu, \quad (\text{VI.9})$$

where c is the velocity and ν the frequency of the radiation, and h is Planck's constant (6.547×10^{-27}). When this energy quantum is equal to or greater than the work function V_0 of the surface, it may eject the electron which absorbs it; the threshold frequency is clearly given by

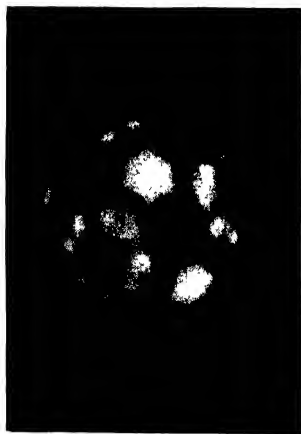
$$V_0 = h\nu_{\min}. \quad (\text{VI.10})$$



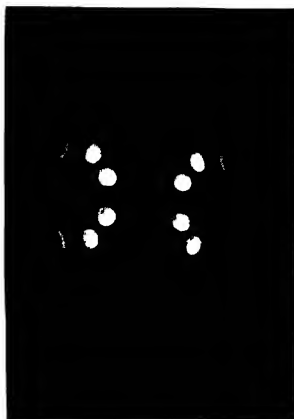
a. Image of etched molybdenum point cathode.



b. Molybdenum cathode after heating to 1520° K. with field applied.



c. Tungsten cathode fully activated with thorium.



d. Thorium on molybdenum cathode, after heating to 1700° K.

PLATE III. POINT-PROJECTOR (OR EMISSION) MICROSCOPE.
(R. O. Jenkins, G.E.C. Ltd., Wembley.)

When ν has values greater than this, the excess energy will appear in the electron, which may be ejected with a velocity up to the maximum given by

$$\frac{1}{2}mv^2 = h\nu - eV_0 = eV_e,$$

V_e being the equivalent emission voltage, and also the counter-voltage which would be required to suppress the emission entirely. It will be remembered that V_0 was defined with respect to a surface at 0° K. At higher temperatures, and especially above $1,000^\circ$ K., the effective work function will be less than this for those electrons which have gained energy above the limiting value W_x (Fig. 84), and therefore a small current will still be obtained for frequencies immediately below ν_{\min} .

Insertion of numerical values in VI.9 gives a practically useful expression for the threshold wave-length λ_0 in Angstrom units:

$$\lambda_0 = 12,400/V_0. \quad (\text{VI.11})$$

For a surface of work function 2 V., therefore, the threshold will lie in the red end of the spectrum ($\sim 6,000$ Å). The photo-electric work function is for some metals identical with that measured for thermionic emission, but it is usually found to be slightly greater. For most metals it lies between 4 and 6 V., so that the threshold falls in the ultra-violet. The alkali and alkaline earth metals, however, have values below 2 V. and give emission of considerable strength in the visible spectrum. Caesium oxidized in a special manner on silver is sensitive far into the infra-red, and gives a maximum output (in the violet) of 60 microamps. per lumen. The inclusion of an inert gas in the cell increases the current, owing to the ionization produced by collision of the accelerated primary electrons with gas molecules. The efficiency of the photo-electric effect (in electrons per photon) is very small, however. Especially sensitive surfaces may yield one electron for each hundred incident photons, but the efficiency is usually much less, and frequently of the order of 1 in 10^4 . The emission increases exponentially with temperature, and is represented by an equation analogous to Richardson's equation (VI.3).

By means of this effect it is possible to produce electron beams of strength proportional to the incident light. The application of a small positive voltage before the sensitive surface, as indicated in Fig. 94, permits of the collection and measurement of the current, after amplification if necessary. If the space is highly evacuated, the output is independent of the applied voltage, above values of a few volts. Simple photo-cells, therefore, find many applications in light photometry, but they can also be adapted for transmitting signals determined by the intensity of an incident beam. When the latter is an optical image it is only necessary to record the photo-current from each elemental area of the surface in turn for television to be in principle achieved; this function is performed by the image tubes discussed in the next section or by the more elaborate 'cameras' described in Chapter VII. Here it remains only to mention that since the threshold wave-length is an upper limit, all radiations of

higher frequency give rise to the photo-effect, including X-rays and gamma-rays. Since, however, only absorption in the surface layer can be effective in ejecting electrons, the yield from these penetrating radiations is even smaller than with visible light.

51. Image Transformer Tubes

Electron optics hardly enters into the design of photo-cells for recording intensities of illumination. It is fundamental, however, to the design

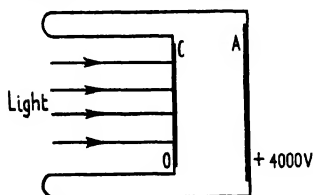


FIG. 95. Holst image transformer.

and operation of tubes of larger size for the conversion of an optical image, projected on to the photo-sensitive surface, into a regular series of electrical signals for transmission to a receiving point (whether by line or radio) where they may be reconverted into a visible image.

This fundamental problem of television may be solved with a mechanical scanning system and a photo-cell of ordinary design, as in the original Baird arrangement. It is more satisfactorily met by an all-electrical method, as was first achieved with the electrostatic and magnetic image converters described in sections 27 and 30 respectively. To these are related the simple image transformers which render visible an infra-red or ultra-violet primary image, or intensify a visible image.

The simplest electrostatic system of this type is that due to Holst, in which the photo-electrons are accelerated directly from the cathode (C) to the neighbouring anode (A) (Fig. 95) by a potential difference of 4,000 V. The anode bears a fluorescent screen, so that an image is formed of much greater intensity than the primary image, owing to the high speed of the incident electrons. The tube is, in effect, a form of electron telescope which increases the brightness rather than the size of the image in order to render it more visible. It may be operated by infra-red or ultra-violet light and thus produce a visible image from invisible radiation.

The electrostatic image-tube of Zworykin (Fig. 45) utilizes the focusing effect of the two-cylinder lens, and the electron image is therefore inverted with respect to the optical image on the cathode. The magnification is again unity, as in the Holst tube. As already mentioned, its aberrations are minimized by employing a high accelerating potential and a curved cathode. The developed form is shown in Fig. 96, where the first cylinder (A_1) is divided into a number of sections carrying increasing voltage, in order to provide variable focusing without impairing the definition of the image; there is a rise of 150 V. between the cathode (C) and the last section, followed by a potential difference of 870 V. between this and the second cylinder (A_2). The electron image is formed on a double-sided mosaic (M) of the storage type (see section 55), which has a multiplier action, and is supported on the signal collecting plate E. Local voltages are established over it depending on the intensity of the illuminating

electrons, and these are scanned off from the farther side by a controlled electron probe (see Chapter VII) in such a way that the electrode E receives impulses proportional to the intensity of each portion of the

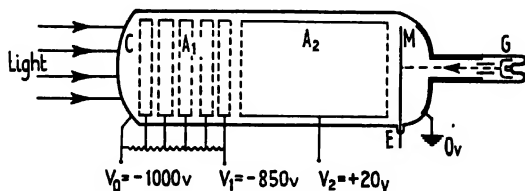


FIG. 96. Image converter with double-sided mosaic (M): the anode system (A_1, A_2) forms an image with photo-electrons on the left-hand side of M , the right-hand side of which is then scanned by a beam from the gun G . (Zworykin.)

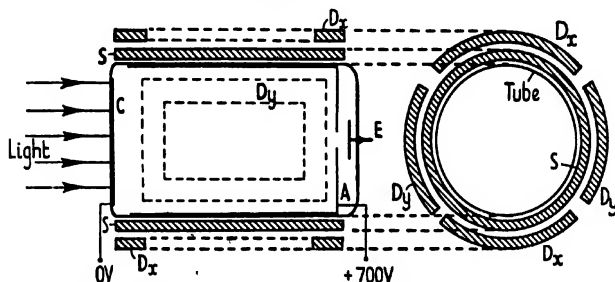


FIG. 97. Farnsworth image dissector: the photo-electrons from C form an image on A which the scanning field due to D_x, D_y then displaces so that each element in turn falls in the aperture before the signal-plate E .

image in turn. These form the television signals for transmission to a distant reconverter. The system proves to be about 10 times as sensitive as the earlier tubes, on account of the multiplying action of the mosaic. The latter is difficult to construct, however, and the single-sided but more complicated arrangements (iconoscope, emitron, and orthicon) described in the next chapter are used in practical television.

The Farnsworth image tube (Fig. 97) is a developed form of an 'image-dissector' proposed by Dieckmann and Hell in 1925. The accelerating voltage is graded across a graphite coating on the wall of the tube between C and A , so that the electron beams from each optical image point are drawn off in fine pencils and so form small circles of confusion in the electron image. True image formation is ensured by the solenoidal magnetic field (S). Each part of the image is then brought in turn into the collecting aperture in the anode by the scanning action of two pairs of magnetic deflecting coils (D_x, D_y) set at right angles. The signal received by the collecting electrode is then proportional to the light intensity in the corresponding point of the optical image. Later versions of the tube incorporate an electron multiplier to amplify the signal before injection into the normal amplifying circuits. However, the primary signal is less than that in the iconoscope, which has accordingly been preferred in practice.

52. Secondary Emission

An incident beam of electrons is capable of ejecting electrons from a metal surface, in much the same way as do photons in the photo-electric cell. The essential requirement again is that the energy of each individual in the beam should be greater than the work function of the surface; this is at once the case if the beam has been accelerated through a greater potential difference than V_0 . It follows that electron beams moving under a potential difference of more than a few volts are adequate to produce such emission: in the case of ion beams the required voltage is less in inverse proportion to the mass of an ion relative to that of the electron. In either case the incident energy may be shared among a number of electrons in the metal so that more than one may be ejected at a single impact.

The emitted, or secondary, electrons are mixed with reflected primary electrons. The relative proportions may be roughly established by velocity analysis, which shows a large low-velocity component, a small high-velocity component due to reflection, and an intermediate straggle of velocities. It is normally not necessary to distinguish between emitted electrons as to origin, and all are termed secondary electrons. From most metals, when carefully cleaned, between 1 and 2 secondaries appear per primary electron, although the ratio is in some cases as high as 4 or 5. Ions are much less efficient, helium ions producing only 1 electron per 5 incident ions. Contaminated surfaces, on the other hand, have a high ratio for electrons, especially when coated with a thin film of rubidium or caesium or their oxides or halides. A Cs-O-Ag surface after special 'forming' by heat treatment may have a secondary emission ratio as high as 10 or 12, for a primary voltage of about 500 V., and may give an output of 60 ma./watt. The activation process for optimum emission is not greatly different from that needed for photo-electric sensitization.

The details of the collision process in the surface are not yet explained. In order to transfer energy and eject a particle in the opposite sense, the incident electron must be involved in a 3-body collision of some sort, in view of the necessity for conservation of momentum. It is possible that it is the bound electrons, and not those in the conduction levels, which are ejected. Support is lent to this view by the fact that insulators as well as metals are found to have a ratio greater than unity.

If the secondary electrons from one surface are directed by a suitable field on to another sensitive surface, the output will be again increased by the emission ratio: if this is 10:1 the output current will now be 100 times the input current. Electron multiplier tubes (next section) operate in this manner; by employing several successive stages, amplifications of the order of 10^6 may be obtained. The other main application of secondary emission is in television pick-up cameras, as already indicated.

The output ratio from a single surface may be raised intermittently to the order of 10^3 by incorporating an insulating layer. The system then

combines both field and secondary emission, and is known as the Malter effect. The essential features are shown in Fig. 98, where M is the metal substrate, I is the thin film of semi-conducting material deposited on it,

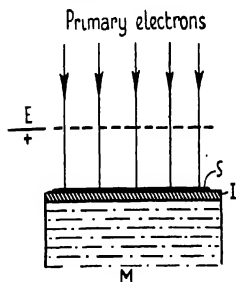


FIG. 98. Malter effect: S is a secondary emitting film on a semi-conducting layer I deposited on a metal base M .

and S is a very thin surface film of high secondary emission ratio. In Malter's experiments S was a caesium film deposited on an oxidized aluminium surface. An electrode E before the surface serves at the same time to accelerate the primary electrons towards it and to collect the emitted secondary electrons. The loss of the latter rapidly raises the potential of the surface towards that of the collector. A high field strength is thereby set up between S and the metal substrate across the thin semi-conducting layer of oxide. The reduction in the work function of M is enough to allow electrons to be extracted, and this cold emission passes through I to the caesium surface and augments the secondary emission. The total emission initially is found to be of the order of 10^3 greater than the primary current, but falls as the surface potential approaches that of the collector and space charge accumulates between them. For this reason the arrangement is best suited to intermittent operation. On cessation of the primary bombardment the emission is observed to continue for some time, until the potential difference across the oxide film becomes too low to maintain field emission. This storage effect may, in favourable circumstances, persist for as long as an hour. It causes the output to be slow in response to changes in input current, and hence limits the practical utility of the effect.

53. Electron Multipliers

In order to utilize the secondary emission effect in amplifying small currents it is only necessary to arrange a series of sensitive surfaces in such a way that the emission from each is directed on to the next. The initial current may be thermionic or photo-electric in origin, and a small fraction of a microampere in value; the output will depend on the number of stages of multiplication, and the gain in each, but will frequently be of the order of milliamperes.

The simplest electrode arrangement would be a series of parallel plates, each staggered with respect to the previous one and a few hundred volts higher in potential; the electrons would then be accelerated from plate to plate by the electrostatic field. A more efficient system was proposed by Slepian in 1923, and later perfected by Zworykin, based upon a combination of electric and magnetic fields very similar to that used by J. J. Thomson for experiments on the nature of photo-electric particles (Fig. 5, p. 8). In the multiplier the continuous plates are split into a number of parallel sections electrically connected in cascade (Fig. 99a).

The photo-electric surface P is at earth potential and light is incident on it through the gauze G parallel to it and 20 to 50 volts higher in potential, being directly connected to the secondary emitting surface S_1 next to P . A transverse magnetic field is established over the whole length of the tube, directed into the plane of the diagram (Fig. 99 *b*) so that the electrons emitted from P describe a path which to a first approximation is cycloidal (see section 7) in the crossed magnetic and electrostatic fields,

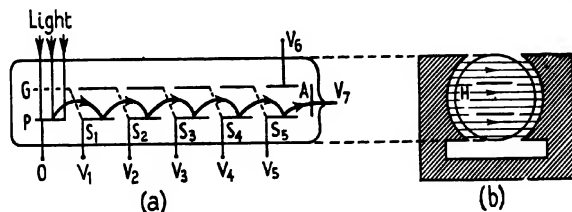


FIG. 99. Slepian electron multiplier: (a) longitudinal section of tube; (b) transverse section of tube and electromagnet.

and then impinge on S_1 . There they release a number of electrons depending on the secondary emission ratio of the surface, and these in turn are guided by the fields on to the next sensitive surface S_2 , and so on to the final collecting anode A . If there are 6 stages with a gain of 10 in each, the total amplification is 10^6 and the operating voltage will be of the order of 500 V. The gain per stage and hence the total amplification may be varied by changing the voltage per stage, since this controls the number of secondary electrons per primary. The maximum sensitivity may be as high as several amps. per lumen, but the actual output will be of the order of milliamps., since the tubes are designed to be operated at very low levels of illumination. The threshold of sensitivity is limited by the leakage current of the system, which arises mainly from thermionic emission from the photocathode; it is of the order of 10^{-7} lumen, which is equivalent to about 10^{-12} amp. of photo-current.

The initial current may be thermionic instead of photo-electric in origin, and the tube then takes the place of a multi-stage amplifier of ordinary type. It has the advantage of low background noise as compared with a thermionic amplifying valve, and is therefore to be preferred for handling very small input currents; it is much more compact than a valve amplifier set. Its main disadvantage lies in the technical difficulty of preparing the sensitive surfaces, which must retain their high secondary emission ratios over a long period. The tendency appears to be to employ a single or two-stage multiplier unit for the first step in amplification only, rather than to rely upon it for the total gain required; it clearly cannot handle the large power usual in the output stage of a valve amplifier.

A great number of modifications of the Slepian tube have been proposed, the most notable being the mesh multiplier of Weiss, in which the electrons travel roughly rectilinear paths through a series of grids at successively

higher potentials; the grids are coated with a secondary emitter, and the whole tube is surrounded by a long solenoid which provides the focusing field. Farnsworth has developed a so-called 'A.-C. multiplier' in which electrons are oscillated to and fro at radio frequency between two secondary emitting surfaces, the current being augmented at each impact

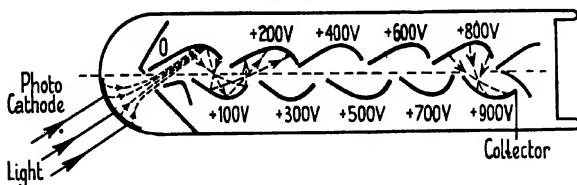


FIG. 100. Electron multiplier with focusing electrodes. (Zworykin and Rajchman.)

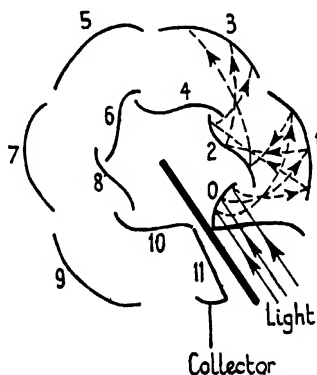


FIG. 101. Circular electron multiplier with focusing electrodes. (Zworykin and Rajchman.)

with a surface. A limit is set to the multiplication by the accumulation of space charge and by the tendency of the system to break into self-oscillation. An overall gain of 2,000 at an output of a few milliamps. has been obtained. Modified versions of it have been proposed both as multiplier tubes and as cold cathode short-wave oscillators.

The multiplier systems described above, and most of their modifications, embody both electrostatic and magnetic focusing. A considerable advance was made by Zworykin and Rajchman, who made a detailed investigation of electron trajectories in purely electrostatic fields of different forms, using the rubber model method (section 19). They found that the most efficient arrangement for an extended tube was that shown in Fig. 100, where each surface is 100 V. higher in potential than the preceding one and has the same shape in all planes parallel to that of the diagram, up to a length of several centimetres. A more compact arrange-

ment was also devised (Fig. 101), in which the sensitive surfaces are arranged around the perimeter of a circle. The latter form is the basis of a commercial multiplier tube which has 9 stages and operates at a maximum voltage of 900 V. Owing to its low noise level it is capable of detecting an intensity of illumination 200 times smaller than the limit for a normal photocell followed by a conventional amplifier.

FURTHER READING

1. Electrons in Metals:

- Crowther (J. A.), *Ions, Electrons and Ionizing Radiations* (1938).
 Wilson (A. H.), *Semi-conductors and Metals* (1939).
 M.I.T., *Applied Electronics* (1943).

2. Thermionic Emission and Electron Guns:

- Richardson (O. W.), *Emission of Electricity from Hot Bodies* (1922).
 References under (1).
 Langmuir (D. B.), *P.I.R.E.* 25, 1945, 1937.
 Pierce, *J. Appl. Phys.* 10, 715, 1939; 11, 548, 1940.
 Law, *P.I.R.E.* 25, 954, 1937.
 Gas Focusing: v. Ardenne, *P.I.R.E.* 20, 1310, 1932.

3. Field Emission; and Projection Electron Microscope:

- Schottky, *Z. Phys.* 14, 63, 1923.
 Fowler and Nordheim, *Proc. Roy. Soc.* 119, 173, 1928.
 Dosse and Müller, *Z. Phys.* 119, 415, 1942.
 Jenkins, *Reports on Progress in Physics*, 9, 177, 1943.
 Slack and Ehrke, *J. Appl. Phys.* 12, 165, 1941.
 Johnson and Shockley, *Phys. Rev.* 49, 436, 1936.

4. Photo-electricity and its applications:

- Campbell and Ritchie, *Photoelectric Cells* (1929).
 Hughes and Du Bridge, *Photoelectricity* (1934).
 Zworykin and Morton, *Television* (1940).
 Farnsworth, *J. Franklin Inst.* 218, 411, 1934.

5. Secondary Emission and Electron Multipliers.

- Zworykin and Morton, loc. cit.
 Farnsworth, loc. cit.
 Weiss, *Z. techn. Phys.* 17, 623, 1936.
 Malter, *Phys. Rev.* 49, 478, 1936; 50, 48, 1936.
 Zworykin and Rajchman, *P.I.R.E.* 27, 558, 1939.

CHAPTER VII

THE CATHODE-RAY TUBE AND ITS DERIVATIVES

54. *The Cathode-ray Tube*

THE essential components of the cathode-ray tube have already been described, either directly or by implication. It remains here only to discuss the instrument as a whole, its uses, and some of its limitations. A diagram of the complete electrode system with its glass envelope is given in Fig. 102. The electron gun, which is shown as a simple two-anode type, provides a fine beam of electrons from an oxide-coated cathode indirectly heated. In place of the anode system a magnetic focusing coil M may be used for focusing the beam on to the fluorescent screen S . The two sets of parallel plates XX' , YY' serve to deflect the beam as required in a horizontal (x -) and vertical (y -) direction respectively. Each pair is arranged symmetrically about the axis, the XX' plates being above and below the plane of the diagram; potentials may be applied to them by leads entering through the sides of the tube (as shown) or through the base, which supports the electrode assembly and carries the connexions to the elements of the electron gun. The interior of the glass tube is normally coated with a conducting layer of graphite, which is at the potential of the final anode and serves to transmit to earth the electrons received by the screen and any others that may be scattered from the beam on to the walls.

In operation the focusing of the beam is controlled by the voltage on the first anode, which is of the order of hundreds of volts positive with respect to the cathode, which in turn is several thousand volts negative to the final earthed anode. The intensity of the beam is varied by changing the small negative voltage applied to the grid. In the magnetically focused type the current in the coil M controls the focus, the only variable voltage being that on the grid. In the gas-focused tube the intensity may be controlled either by a grid or by the overall voltage applied.

When a positive voltage, say, is applied to one of the deflecting plates the beam will be attracted towards it and at the same time accelerated. The increase in speed may be enough to put the spot out of focus on the screen, at maximum deflexion. This defocusing effect may be avoided by applying an equal negative voltage to the opposite deflecting plate; this is called symmetrical deflexion. For many purposes, however, it is sufficient to earth one plate of each pair and to apply the deflecting voltages to the other plates only. When a voltage is applied across each pair in turn, the beam will be deflected in the x - or y -direction as the case may be, and the visible spot will take up a new position on the screen. If the voltage is alternating, then a horizontal or vertical line will be observed as the beam sweeps to and fro across the screen, owing to the

persistence of fluorescence in the material composing the latter. When the two deflecting fields are applied simultaneously, the spot will describe a complex path depending on the relative amplitude, phase, and frequency of the applied potentials; such figures are known as Lissajous figures. They offer a means of comparing an unknown with a known alternating supply.

The deflexion of the beam may be produced magnetically as well as electrostatically (sections 3 and 4), but since the magnetic force is normal both to the axis of a coil and to the initial direction of the electron, the two pairs of deflecting coils are displaced through 90° with respect to the corresponding X - and Y -deflecting plates. The deflecting influence is now due to the current in these coils, and not to an applied potential.

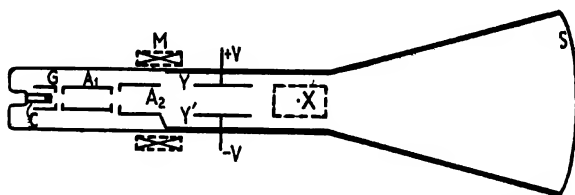


FIG. 102. Cathode ray tube with electrostatic focusing.

Magnetic deflexion therefore has the advantage of not changing the velocity of the beam, and hence of avoiding defocusing. On the other hand, the voltage needed to produce a given current in the coils, and thus a given deflexion, depends on the frequency employed. Where a wide variation of frequency is involved, electrostatic deflexion is to be preferred. The deflexion sensitivity is discussed in section 56.

The screen S is coated with a fluorescent material giving a bright response to the incident electron beam, such as zinc sulphide or phosphate, calcium tungstate, or a mixture of these or other substances. The duration of the fluorescence may vary considerably, and the material used is chosen with regard to the nature of the impulses to be applied to the tube. If they are of high frequency giving a rapidly moving image, then a short-period fluorescence will be demanded; but if transient phenomena are to be investigated, it will be advantageous to use a screen material with a slow decay, so that the trace persists for some time. Similarly the type of screen will depend on whether observation is to be visual or photographic. In television the colour of the response is important, and almost any colour can be produced with the screen materials now available. These are not problems of electron optics, properly speaking, and will not be discussed here; a great deal of information is available in the technical literature.

The cathode-ray tube finds application in very many fields where periodic phenomena are to be studied. Almost any simple effect may be translated into a transient or oscillatory voltage and applied to the deflect-

ing system of a calibrated tube. This is especially true of radio and television experiments, but a great variety of other problems are suitable for investigation with its use, from the mode of action of muscle to that of an internal-combustion engine. The technical problem is only that of devising a suitable means of producing or picking up a voltage change, and then of amplifying it to a magnitude sufficient to produce a significant deflexion of the electron beam. According as the tube is used for direct observation or for photographic recording, it may be known as an oscilloscope or oscillograph. Details of particular applications will not be given here, since many excellent books on the subject are now available.

The Double-beam Tube. For many purposes, both in research and routine work, it is necessary to compare two phenomena as directly as possible; they may be independently produced or may result from the operation of two variables on the same system. They may be two frequencies, the variation of voltage and current in the same circuit, or a primary impulse and its echo, for example. By the use of an electronic switching circuit the two impulses may be applied alternately to a cathode-ray tube at such a rate that the resulting patterns appear continuous, owing to the persistence of the effect on the screen and the flicker limitation of the eye. It is simpler, however, to employ a tube with a double beam and two deflexion systems. In early models the beam was divided into two portions by inserting in the anode of the electron gun a diaphragm in which two holes were bored. The method now used is to fix a splitting-plate edge-on to the beam where it leaves the final anode of the gun. The plate extends between the *Y*-plates, dividing them into two separate deflexion systems; it is connected directly to the final anode and is thus at earth potential. Each *Y*-plate then affects one half of the beam independently, giving two vertical traces which are spatially π out of phase when the deflecting impulses are in phase. The *X*-plates are common to both beams, giving the normal horizontal deflexion; there is therefore absolute simultaneity between the two traces on the screen. The Cossor double-beam tube incorporates a system of this type; it may also be operated as a single-beam tube when desired. Cathode-ray tubes with three and more beams have also been described.

The Kinescope. In television the cathode-ray tube acts as reproducing system for an image transmitted by radio waves. The beam in this case not only has to be deflected in a regular manner to all parts of the screen in turn, but also has to be modulated in intensity to an extent proportional to the light intensity of the original object. The deflecting or scanning system already described fulfils the first requirement, but for modulation an extra electrode may be inserted in the tube. When used for reproduction in this special way, the tube is often called a kinescope. In principle, the modulating impulse may be applied to any of the electrodes of the assembly. However, a pulse of given size will have much less effect on the beam when applied to an electrode at high potential than to one at low potential, and it will be most effective when applied in a region where

the velocity of the electrons is low. Hence it is normally imposed upon a grid at negative potential, an additional advantage being that no current is then taken by the modulating electrode. The simple grid of the three-electrode gun may be utilized, or a second ('control') grid may be inserted (CG, Fig. 103). In the latter case the first grid may be connected to the cathode, at zero potential, with the result that the emission is considerably increased. The actual current in the beam in such tubes ranges from 100 μ a. up to many milliamperes, and the total grid variation is of the

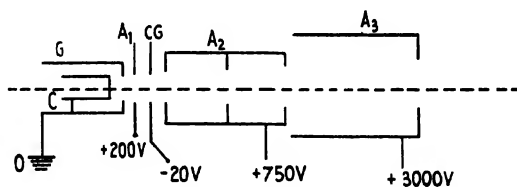


Fig. 103. Three-anode electron gun with modulating grid.

order of 20 V. The final anode voltage may be 5,000–7,000 V., and the first anode at about 1,000 V. An additional anode is frequently employed to accelerate the beam rapidly from the cathode; it may be at less than the potential of the second anode, when it will be placed close to the grid (A, Fig. 103), or the system may be a symmetrical lens, in which case it will have the same high potential as the final anode.

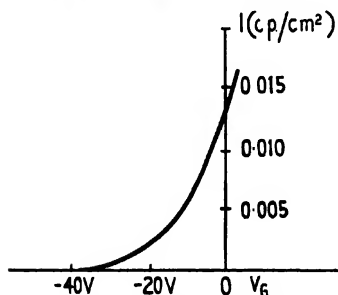


Fig. 104. Modulation characteristic for medium voltage cathode-ray tube. I = intensity of illumination. V_g = control grid voltage. (Zworykin and Morton.)

The modulation characteristic for such a medium-voltage tube is shown in Fig. 104, grid voltage being plotted against the intensity of illumination (I) in the image spot (in candle-power per cm.²). It has the general form of the lower bend of an ordinary triode valve, the slope increasing rapidly with increase of voltage, until a roughly constant slope is obtained; in terms of variation in the exciting beam current, this slope is of the order of 30 ma./V.

When the image on the screen has to be projected optically to a large size for communal viewing, its brightness has to be increased as much as possible. For this purpose a beam of wide initial aperture must be employed to increase the total current, and the overall accelerating voltage raised so that the electrons transfer high energy to the screen, which must be of high luminous efficiency. The voltage on the final anode may therefore be as high as 100,000 V., in large tubes. A wide beam is best collected by the magnetically focused tube, described in section 47 (Fig. 90), which has a large aperture. Alternatively,

the Law type of gun may be used (Fig. 91), in which the aperture is small and a cross-over of high intensity is obtained in a gun of special design.

In circumstances where great deflexion sensitivity is required, but nevertheless the final beam velocity must be high in order to give sufficient brightness on the screen, an additional anode may be introduced to give post-deflexional acceleration. In one type the tube is fitted with a narrow graphite coating in the form of a ring around the wall, immediately before the screen. The voltage applied to the last anode of the gun proper is comparatively low, and the graphite ring is maintained at twice this potential. The beam is thus deflected when it is at intermediate velocity, and the deflexion sensitivity is correspondingly high (cf. section 56). It is then accelerated to a high speed by the field between the ring coating and the gun anode, and produces satisfactory brightness in the image on the screen.

55. *Television Transmitting Tubes*

In a television system the signals which actuate the kinescope and produce a visible image at the receiving end are generated in the first place in a special type of camera in which a light optical image is converted, element by element, into a series of electrical impulses. The image, formed on a special surface, produces a temporary electrical (or, rather, photo-electrical) effect instead of the chemical action which, occurring in a sensitive photographic plate, results in a permanent representation. It is usual for the electron image to be scanned, and neutralized, by an electron beam. Such a camera is variously known as an iconoscope (the R.C.A. trade-mark), as an emitron or super-emitron (in Great Britain), and as an orthicon, in a new and more sensitive form. In the Farnsworth dissector tube, on the other hand, no scanning beam is used, but the electron optical image as a whole is periodically displaced so that the current density in each part is received in turn on a special collecting electrode. This type has already been described in connexion with image transformers (section 51, Fig. 97). The iconoscope type is more closely related to the cathode-ray tube, however, and is conveniently treated here, so far as its electron optical construction is concerned. The general features of the transmitting and receiving equipment, and the means by which the one is controlled by the other, are dealt with later in this section.

The Iconoscope. The essential components of the camera are a multi-element photo-sensitive screen ('mosaic'), an electron gun and deflecting system for scanning it with an electron beam, and a signal plate which acquires a voltage proportional to the brightness of the element of the image momentarily under the scanning beam. The simplest arrangement is that of the original iconoscope, shown in Fig. 105, where the light image is focused directly on the mosaic *M*, behind which is the signal plate *S*; the electron gun *E* is disposed in a side-arm below the mosaic. The normal potential of the mosaic is the same as that of the final anode of the gun *A*, which is a graphite coating on the wall of the tube and

serves also to collect electrons emitted from the mosaic. S is slightly negative with respect to the mosaic.

The operation of this system requires that the mosaic consist of a number of minute conducting elements, well insulated from each other, so that the photo-electric image is composed of independent points carrying a charge determined by the local intensity of illumination. This condition is met by depositing a very thin film of silver on a mica sheet, which is then heated to 700°C . At this temperature, well below the melting-point of silver in bulk, the metal atoms aggregate together into globules under

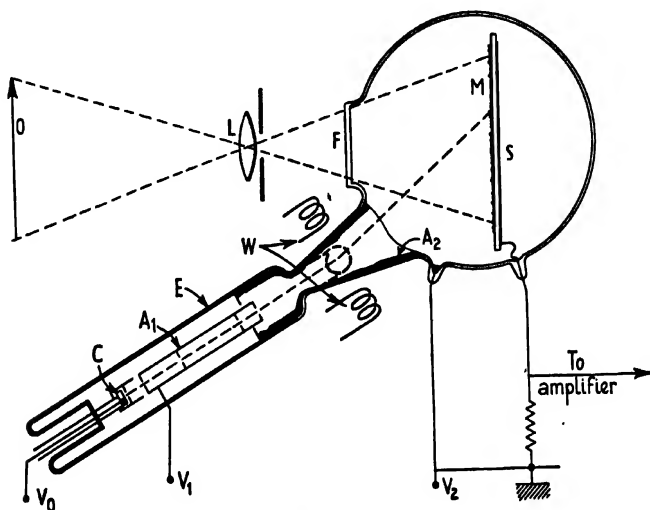


Fig. 105. Iconoscope type of television camera.

surface-tension forces. Under favourable circumstances the globules may be as small as a few microns in diameter, as microphotographs show. They are then sensitized with caesium, and thus form a multitude of insulated photo-cells, which are also secondary-emissive under electron bombardment, having an emission ratio near to unity under the beam voltage employed. The mosaic screen is usually a square of about 10 cm. side. The signal plate S is in the form of a metallized coating on the reverse side of the mica; with respect to it each photo-cell forms a minute condenser. The discharge of any element will result in a corresponding impulse in the signal plate. Between each periodical discharge, whilst the beam is scanning all the other elements, an individual globule will be continuously emitting photo-electrons and acquiring charge. The mosaic thus exhibits a storage effect, and its sensitivity is greater than if each element were illuminated in turn, or called into use once per picture, as in the Farnsworth tube. The requirement follows that the elements must be well insulated from each other, so that the accumulating charge is not dissipated

during a scanning interval. Such a mosaic was first suggested by Campbell Swinton in 1908 and, when modern technique offered the possibility of its realization, he carried out experiments on it in 1926. The first practicable system, however, was described by Zworykin in 1933, the patents having been taken out several years earlier; it was essentially the same as the tube described here.

The electron gun is of simple form, consisting of a cathode C , grid G , and two anodes A_1 and A_2 , the latter being the graphite coating. It produces a beam of moderate velocity, the overall potential usually lying in the range 100–1,000 V. In the undeflected condition it is focused at the centre of the mosaic, or somewhat below it. To keep it in focus over the rest of the screen during the scanning process, an auxiliary electrode is sometimes used, modulated so that the beam is automatically changed in focus. Usually, however, the beam current needed for scanning is so small ($\sim 0.1 \mu\text{a.}$) that the gun aperture can be reduced to such a value that the depth of focus is great enough for the beam to remain sharp over the whole screen. The horizontal and vertical deflexion of the beam is caused by two pairs of magnetic coils W , arranged around the neck of the side tube. In order to achieve high definition in the transmitted image the mosaic has to be dissected by the scanning beam into a great number of elements, in the form of fine horizontal lines; in American usage 441, in the B.B.C. system 405 lines. The primary requirement of the electron gun, therefore, is the production of a beam to throw on the screen a spot of the order of 0.1 mm. in diameter, which will remain of essentially the same size during deflexion through a wide angle. Fortunately this is not difficult to achieve with a gun of normal design when the beam current is as small as is required here. Especial attention, however, has to be paid to the elimination from the beam of scattered and secondary electrons produced from the gun elements, by proper disposition of the limiting apertures in the anode.

In operation the tube functions as follows, making the assumption of a perfect mosaic so that leakage of charge from one element to another is excluded. A light image of an external object or scene O is projected by the lens system L through the optically flat surface F at the front of the tube, on to the mosaic M . Photo-electrons will be emitted from each globule in proportion to the intensity of the light falling on it, until such time as the electron beam passes across and discharges it, after which photo-emission will begin again. The beam scans each element 25 times per second (in the B.B.C. system) and 405 lines are covered in each scanning period; it therefore remains on each globule for less than one-millionth of a second. The remainder of each period is available for emission of photo-electrons, under constant illumination, and hence the storage effect should considerably increase the efficiency, in terms of signal strength per lumen.

After scanning, each globule is found to be at about -1.5 volts with respect to the mica substrate. Those which are unilluminated

remain at this potential until the beam returns again; under its impact they emit secondary electrons and rise in potential to about $+3$ V., at which potential the secondary emission ratio has fallen to unity. After the beam has passed, some of the secondary electrons emitted by the rest of the mosaic fall upon them and reduce them again to about -1.5 V. The scanning of each darkened element therefore changes by 4.5 V. the potential across the condenser formed by the element and the signal plate, resulting in a minute current flow through the load resistance, which is of the order of $10,000$ ohms. An illuminated globule, on the other hand, will have acquired charge during the frame period on account

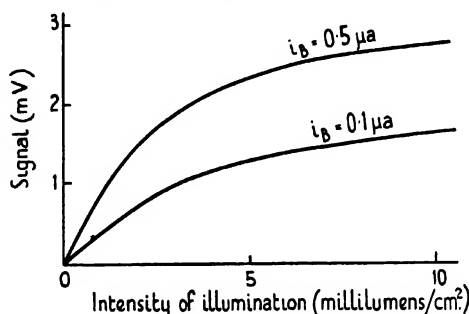


FIG. 106. Iconoscope: signal output with varying illumination for two values of beam current.

of photo-emission, and will rise in potential to some point between -1.5 and $+3$ V., depending on the brightness at this point. The scanning beam can only raise it to the limiting potential of 3 V., and hence the signal transmitted to the plate is less than 4.5 V. in proportion to the local intensity of illumination. As the beam scans the whole picture, a train of impulses will be passed to the signal plate and into a radio amplifier for transmission.

The variation in output with illumination is roughly linear over an appreciable range. In Fig. 106 is shown the output (in millivolts) across a $30,000$ ohms input resistance to the amplifier for two values of the beam current. It will be seen that the higher the illumination the more the characteristic departs from linearity. The other main disadvantage of the simple iconoscope is the small efficiency (5 – 10 per cent.) of photo-emission, owing to the mosaic being almost at the potential of the collecting electrode and hence the accelerating potential for the photo-electrons being very small. For the same reason stray secondary electrons fall back on to the surface and interfere with the storage action in the photo-elements. The output proves to be only one-half that of a similar photo-surface in continuous operation, indicating that the storage effect operates to a small extent only.

The most direct method of increasing the output is to incorporate a secondary emitting surface to which the primary photo-electrons are

accelerated, thus separating the photo- and scanning processes. This was achieved by Zworykin in the image tube already described (section 51, Fig. 96), in which the electron image is formed on one side of a thin mosaic, the other side of which is scanned by an electron beam from a gun situated on the axis. This tube has a sensitivity about 10 times that of the iconoscope and the advantage of straightforward construction. The double-sided mosaic proved difficult to prepare, however. The method used was to cover the wires of a fine nickel gauze with an insulating coating of vitreous enamel. Silver oxide was then pressed into the interstices and reduced to metallic silver by heating *in vacuo*.

The Super-Emitron. The earlier super-emitter of Lubszynski and Rodda

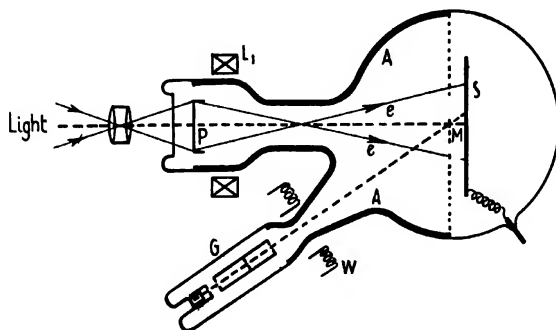


FIG. 107. Super-emitter television camera. (McGee and Lubszynski.)

avoided the difficulty of making a double-sided mosaic by using a beam to scan the screen from the same side as the incident illumination (Fig. 107). The optical image ejects photo-electrons from the transparent cathode *P*, which are all drawn off by the high potential (400 V.) maintained between *P* and the anode coating *A*. Owing to the absence of secondary emission in this stage, no stray electrons fall back on the photo-elements to reduce the emission. The photo-electrons are focused on to the mosaic by the magnetic lens *L*₁. The electron optical system produces a linear magnification of 4:1 in the image, which is also rotated by the action of the magnetic lens. As before, the signal plate is a few volts negative to the caesium-silver mosaic, with a sheet of mica 0.25 mm. thick separating them. The electron gun *G* provides a scanning beam at 400 V., the final focusing being effected by the collector surface coating; the deflecting system is magnetic. Owing to the increased efficiency of photo-emission and the multiplying effect at the mosaic, the sensitivity of the tube is found to be 10–15 times that of the earlier emitter, which was of the simple iconoscope type and gave 12 μ a. per lumen. The super-emitter was used in the B.B.C. transmissions from the Alexandra Palace up to the commencement of war in 1939.

The Orthicon. The orthicon of Iams and Rose, of slightly later date, is

of essentially similar construction, but operates with a low-voltage scanning beam. The purpose is again to ensure that no secondary electrons are emitted from the scanned surface, which may be the primary photo-electric surface or the reverse side of a secondary emitting mosaic; at the same time the photo-electrons are accelerated off the primary surface by a strong field, so that maximum emission is ensured. The scanned surface is normally at the same potential as the cathode of the electron gun, and the beam is accelerated by 25 V. only. During the scan it discharges each illuminated globule in turn, returning to it just as many electrons as were

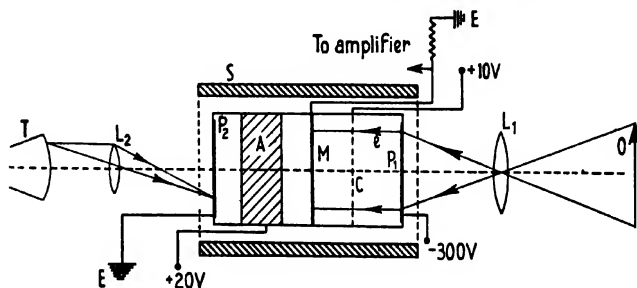


FIG. 108. Orthicon television camera. (Iams and Rose.)

lost by photo-emission, and thus bringing the whole surface to cathode potential again; the beam is repelled from unilluminated globules, which thus take no current and generate no signal. The corresponding current to the signal plate, passing through a resistance, gives an output signal in direct proportion to the initial illumination of each globule: hence the name of orthicon.

In order to achieve true linearity of response, however, the scanning beam must strike the elements at normal incidence and a special deflecting system is required to ensure that this occurs at all parts of the surface. Ingenious use is made of photo-electric or thermionic beams, no electron gun in the normal sense of the term being employed. One experimental arrangement is shown in Fig. 108, where light is incident from the right on a transparent photo-cathode P_1 ; electrons emitted from its further side are accelerated on to a double-sided mosaic M under a potential difference of 300 V. The secondary electrons emitted from the image on M are collected by the gauze C , at +10 V. The reverse side of M is scanned by a beam of electrons from the scanning photo-cathode P_2 , where they are produced by the moving image of the spot on the screen of a cathode-ray tube T , projected by the lens L_2 ; they are accelerated on to M by the wall-electrode A , and maintained as a beam by the action of the longitudinal field of the surrounding solenoid SS . Beam currents up to $2 \mu\text{a.}$ are produced by this photo-electric method. The system was found to operate satisfactorily, although the resolution was not so high as in the iconoscope or emitron; the tube described was only in the development stage, however.

The saturation of photo-emission ensures a high sensitivity, which is claimed to be as high as that of the multiplier iconoscope and super-emitron. An alternative version gave greater resolution, using a thermionically produced beam deflected in superimposed electric and magnetic fields of special design; a single-sided transparent mosaic was used in this case. The method offers the possibility of improvement in sensitivity, if not in resolution, and has been used in television transmission in the U.S.A.

Television. As television is a comparatively new technique, it may be of interest to conclude this section with a brief account of the processes

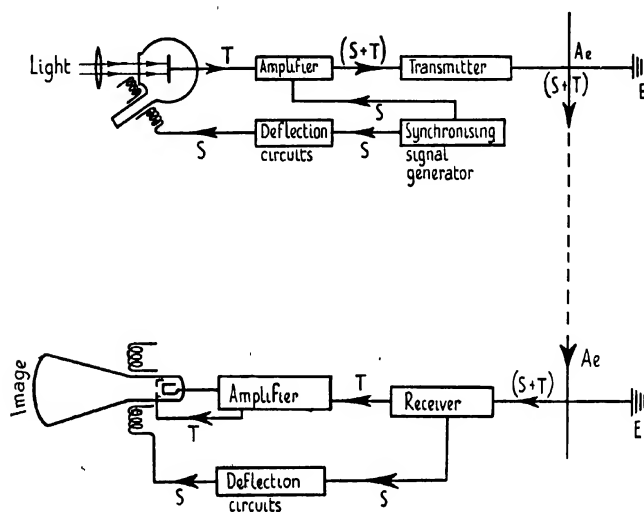


FIG. 109. Television transmission and reception (schematic).

S = synchronizing signal; T = television signal.

involved. The scene to be transmitted is brightly illuminated and focused by a lens system so that a clear image falls on the photo-sensitive surface of the camera (Fig. 109). This image, or the secondary mosaic on which it forms an electron image, is scanned in a regular manner by an electron beam, which makes some 400 horizontal excursions in one complete scan of the image; 25 or 30 such pictures (or 'frames') are taken in each second, the beam automatically returning to the top of the mosaic after each scan. The deflecting system, therefore, has to be supplied with suitable saw-tooth shaped impulses so that the beam is moved at uniform speed along each line, and caused to fly back sharply from the end of one to the start of the next, by the action of the X -plates; the Y -plates cause a similar, but slower, motion from top to bottom of the image followed by a rapid fly-back to the top. The necessary circuit problems will not be considered here, but it is clear that the deflecting system must be of great accuracy (see next section). It is also essential to coherent

reproduction that the scanning process shall occur at exactly the same rate and in precisely the same sequence in both camera and receiver. This is achieved by applying the same controlling, or 'synchronizing', train of impulses to the deflexion system of each tube: directly to the camera and, by incorporation with the television signals, to the receiving cathode-ray tube. The path of the synchronizing signal S is shown in Fig. 109.

The scanning of the mosaic results in a train of television signals T being communicated to the signal plate, their amplitude being (ideally) directly proportional to the intensity of illumination of the corresponding image points. They are amplified and transmitted by short waves to the receiving aerial. The radio receiver detects and amplifies the weak input signals, applying the synchronizing impulses to the deflecting coils or plates of the cathode-ray tube, and the television signals between the grid and cathode of its electron gun. The electron beam is deflected over the whole viewing screen, and its intensity varied according to the strength of the original light signal, in exact synchronization with the television camera. A visible picture is thus produced, the brightness and definition of which will depend on the sensitivity of the camera mosaic and fluorescent viewing screen and on the perfection of the electronic apparatus. In fact the faithfulness of reproduction is remarkably good, taking into account the very short period in which the technique has been developed.

56. *The Deflexion System*

It is convenient to discuss the details of the deflecting or scanning system after describing both the cathode-ray tube and the television camera, since in essentials it is common to both. Many applications of the cathode-ray tube do not demand periodic scanning of the screen by the electron beam, but one or both parts of the deflecting system are always required, if only for static or unidirectional measurements. In television practice, on the other hand, scanning has to occur at high speed and with great accuracy, and the deflexion system must be highly perfected. We shall discuss here the operation, sensitivity, and possible defects of both electrostatic and magnetic methods. In early practice the former was preferred, but increasingly in recent years the magnetic system has been used in both television and ordinary cathode-ray tubes, on account of its simplicity and flexibility, and its greater deflexion sensitivity at high beam voltages.

(a) *Electrostatic Deflexion.* The deflexion produced in a beam when it passes through a transverse electric field has already been mentioned in Chapter I (p. 5). The electrons are caused to follow parabolic paths and are therefore deviated through a small angle during passage between the charged plates producing the field. The system acts as a prism does on light rays, the deviation depending on the velocity or equivalent wavelength of the entrant electrons. If their speed is too low, however, they may be attracted on to the positive plate and thereby eliminated. The amount of deviation produced in a given field can be readily calculated,

under certain simplifying assumptions the validity of which will be examined later.

Let the plates producing deflexion in the y -direction be of length l in the direction z of the beam and of separation d . Potentials V_y , V'_y are applied to them so that the upper plate is at higher potential ($V_y > V'_y$). The field between such plane parallel plates will be essentially normal to their surfaces except at the edges, as shown in Fig. 110. For the present we shall assume that this 'fringing' effect has a negligible effect upon the beam. It will also be assumed that the length of the plates is much less than the distance L to the screen, and that they are sufficiently far from the last lens of the electron gun for there to be no mutual interference

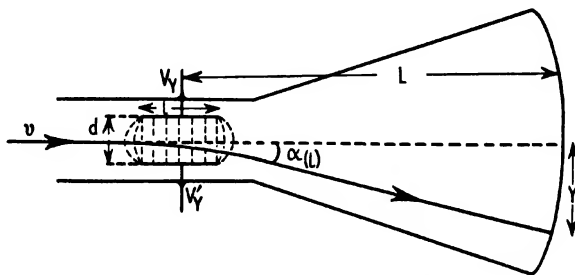


FIG. 110. Electrostatic deflection of beam.

between the two fields. The incident electron beam is supposed to have been accelerated to a velocity v under the influence of the total potential V_0 applied to the electron gun.

Then the motion of an electron will be a straight line on either side of the plates, and within them will obey the equation

$$m\ddot{y} = eE = e(V_y - V'_y)/l, \quad (\text{VII.1})$$

the field E being the negative potential gradient and the acceleration accordingly directed downwards. Taking the origin in the near bounding plane of the system, we integrate and obtain the rate of deviation of the electron:

$$dy/dt = (e/m)Et, \quad (\text{VII.2})$$

the constant of integration vanishing since $dy/dt = 0$ when $t = 0$. As the axial electron velocity may be taken as constant within the field, we may write $v = dz/dt = z/t$, where z is the distance travelled from the origin in time t , and obtain the angle of deviation at any point:

$$\tan \alpha_0 = dy/dz = (e/m)E(z/v^2) = Ez/2V_0, \quad (\text{VII.3})$$

since

$$eV_0 = \frac{1}{2}mv^2.$$

The total deviation suffered by the beam will be given by setting $z = l$, the path length within them, and thus

$$\tan \alpha_0 = El/2V_0 = (V_y - V'_y)l/2V_0d. \quad (\text{VII.4})$$

The actual displacement of the spot on the screen will be

$$Y = L \tan \alpha_0 = (V_y - V'_y) l L / 2 V_0 d \quad (\text{VII.5})$$

and the deflexion sensitivity in cm./volt:

$$Y / (V_y - V'_y) = l L / 2 V_0 d. \quad (\text{VII.6})$$

The sensitivity thus depends upon the geometrical factors as would be expected: it increases with length of plate and distance from the screen, and decreases with separation of the plates. It also proves to depend inversely on the accelerating voltage of the electron beam, and therefore high deflecting potentials are needed in high-voltage tubes.

The fringe fields of the plates are not negligible, however, in their effect on the beam. They form an electrostatic cylindrical lens of power varying with distance from the axis, so that the shape and size of the spot on the screen depend on the degree of deflexion in the y -direction. The effect may also vary with displacement in the x -direction, and thus give rise to a mutual interaction, or 'cross-talk', between the two deflexions. In terms of equation VII.6, the sensitivity will no longer be constant for all values of the deflecting potential. To overcome this difficulty the plates have to be specially shaped and positioned in the tube. Apart altogether from this, they are frequently splayed out towards the screen in order to accommodate a larger deflexion without increasing the separation of the rest of the plate surfaces.

The same considerations apply to the pair of plates causing deflexion in the x -direction. Ideally they could be disposed on either side of the same space as that enclosed by the Y -plates, but it is then found to be difficult to design a shape of plate giving no distortion of the scanning pattern on the screen. In practice the two sets of plates are placed at different distances along the axis. At low frequencies of deflexion, however, this will give rise to a phase difference in the two displacements of the beam, since one will occur slightly later than the other, with respect to a particular electron. The effect can be minimized by increasing the accelerating voltage and thus decreasing the time of passage of the beam between the plates. In special tubes for low-frequency work it may be eliminated by having two pairs of X -plates, one on either side of the Y -plates. As already mentioned, the deflecting voltage is usually applied symmetrically across a pair of plates, and not *in toto* to one plate, in order to avoid influencing the speed of the beam.

(b) *Magnetic Deflexion.* In a uniform transverse magnetic field an electron will describe an arc of a circle (section 3 and Fig. 2), instead of the parabolic path travelled between electrostatic deflecting plates. Therefore magnetic coils have a different deflexion sensitivity from that just evaluated. The deviation produced is also in a different plane, being at right angles to the direction of the field and not in its direction. Hence magnetic deflecting coils of the normal type are shifted through 90° about the axis as compared with the corresponding electrostatic plates. The coils may be simple windings, or they may be provided with an iron yoke

when large deflecting fields are required. For the present purpose it is only necessary to assume that the transverse field produced in, say, the x -direction (H) is uniform over the whole of its extension l along the axis, and is zero outside this space (Fig. 111); the effect of the fringe field will be considered later.

The equation of motion of the electron in the y -direction will then be

$$m\ddot{y} = Hev,$$

and the resulting transverse velocity is

$$\dot{y} = (Hev/m)t. \quad (\text{VII.7})$$

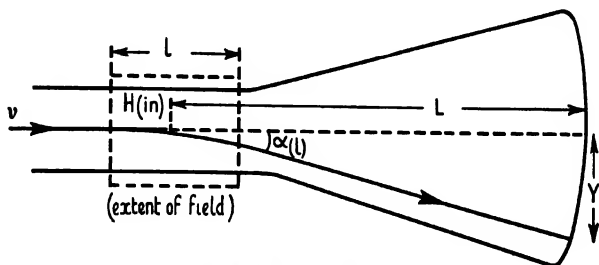


FIG. 111. Magnetic deflexion of beam.

On eliminating time and substituting for v as before, we have for the deviation

$$dy/dz = \tan \alpha_{(z)} = zHe/mv = zH \sqrt{\frac{e}{2mV_0}},$$

the total effect being found by writing $z = l$. The displacement on the screen is then

$$Y = L \tan \alpha_l = HlL \sqrt{\frac{e}{2mV_0}}, \quad (\text{VII.8})$$

and the deflexion sensitivity

$$Y/H = lL \sqrt{\frac{e}{2mV_0}}. \quad (\text{VII.9})$$

For magnetic deflection, therefore, the sensitivity is inversely proportional to the root of the beam voltage, and not inversely to the voltage as in the electrostatic case (VII.6). For this reason the former is preferred for all high-voltage purposes, as in high-speed cathode-ray oscillographs and modern television tubes. Owing to the difficulty of designing a magnetic circuit of low power loss at high frequencies, magnetic deflecting coils were for long used only for low-frequency purposes, including the Y -deflexion of the scanning beam in television. This problem has now been solved, however, and they are now used for the horizontal line scan as well; a typical iron yoke for such coils is shown in Fig. 112a.

The magnetic method has also the advantage that the two deflecting fields can be applied over the same region, with a consequent shortening of

the length of the tube. No problem of the shaping of plates to accommodate the beam arises, nor of interference with the speed of the beam. For these reasons also the tube incorporating both magnetic focusing and magnetic

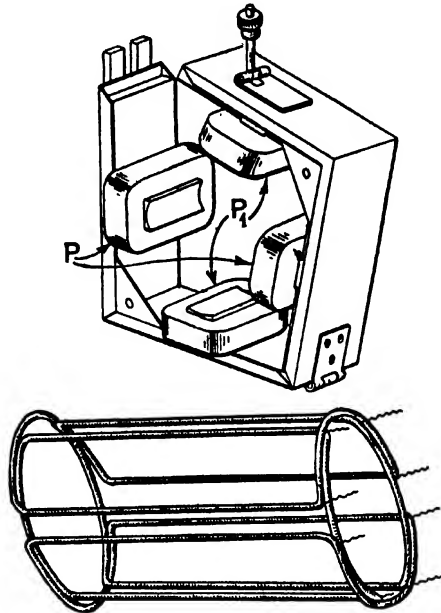


FIG. 112. Magnetic deflecting coils: (a) with iron yoke; (b) without yoke.

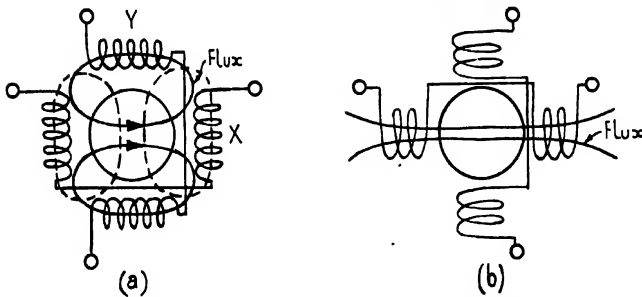


FIG. 113. Magnetic deflecting coils: alternative arrangement.

deflexion is increasingly used. When strong deflecting fields are not required, the coils may be air-cored. They are then of rectangular form, curved to fit the neck of the tube, and are frequently lapped over each other, as indicated in Fig. 112 *b*. The magnetic focusing field may then be superimposed on them. An alternative form of winding, usually upon an iron core, is indicated in Fig. 113; the two coils of a pair are here arranged parallel on either side of the tube.

The fringing field, which was neglected in the above treatment, must now be taken into account. As in the electrostatic case, the field has no abrupt boundary and the resulting lack of uniformity at either end of the deflecting system causes aberration of the beam; it may also distort the scanning pattern. The actual distribution of the field may be as shown in Fig. 114*a*, in longitudinal and transverse cross-section through the iron poles. The departure from uniformity has been exaggerated in order to

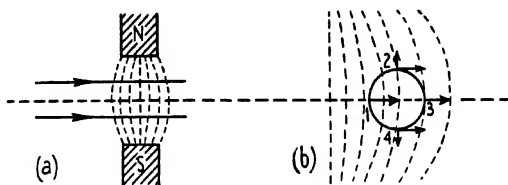


FIG. 114. Distortion in magnetic deflecting field (exaggerated). (a) Longitudinal section, (b) enlarged vertical section, of initially circular beam. (After Zworykin and Morton.)

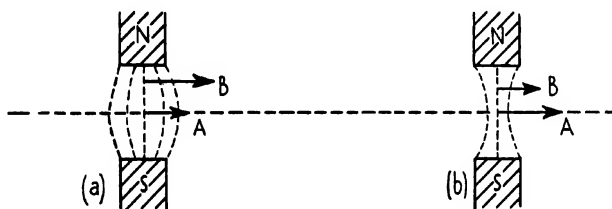


FIG. 115. Distortion in magnetic deflecting field (exaggerated): (a) barrel field; (b) pincushion field. (After Zworykin and Morton.)

make the effects clear. If the beam is originally circular in cross-section (Fig. 114*b*), then rays 1 and 3 on a horizontal diameter will suffer different displacements, as indicated by the force vectors. The vertical field is stronger at 1 than at 3, and therefore in being deflected towards the right the beam will also be contracted in this sense. On the other hand, rays 2 and 4 on a vertical diameter through the beam experience the same horizontal force but vertical forces in opposite senses, since the horizontal component of the magnetic field is finite and oppositely directed at the two points. Hence the beam will suffer an elongation in the vertical plane, and the resulting shape of the deflected spot will be elliptical. If the beam is not in the horizontal mid-plane of the field, as assumed here, the ellipse will be rotated about the axis also.

Distortion of the scanning pattern arises from variation of field strength with distance from the axis, which may be of barrel or pincushion form (Fig. 115). In the former case (a) the beam will enter a region of higher field strength as it moves away from the axis, and thus experience a greater deviation, as indicated by the longer force vector at *B* than at *A*. The outer parts of the scanning pattern will experience a greater relative

displacement than the axial portion, and the image will exhibit pincushion distortion. Conversely, when the field shows pincushion distortion (Fig. 115*b*), the beam will be in a stronger field near the axis (*A*) than away from it (*B*), and the scanning pattern will suffer from barrel distortion. The remedy lies in careful design of the windings and magnetic circuit (if any) of the deflecting coils. The finished system of a high resolution television tube is of some complexity.

In the above discussion of deflexion problems it has been assumed that the beam follows a scanning pattern of equally spaced horizontal lines of equal length. For special purposes it may be desirable to have a circular trace on the screen, upon which incoming signals can then be imposed. No electron optical problem is involved in effecting this, it being only necessary to apply to the *X*- and *Y*-plates equal deflecting impulses of sinusoidal form; if a phase difference of $\frac{1}{2}\pi$ is maintained between the two alternating potentials, the resulting deflexion of the beam will produce a circular path on the screen.

57. *High-Speed Oscillographs*

A further complication affects the deflexion system at very high frequencies of operation, when the time of transit of an electron through the field may be comparable with the period of the deflecting impulse. Electric and magnetic systems will be affected equally. The speed of the electron beam at an accelerating potential of a few thousand volts is of the order of 10^9 cm./sec. A given electron will thus take about 10^{-3} microsecond to pass through the deflecting system. If the latter is subjected to an alternating potential of frequency 1,000 Mc./sec., then one complete wave impulse will occur during the time the electron is between the plates. It will, therefore, receive no resultant deflexion at all, since the integral of the impulse over the complete cycle is zero, and the same state of affairs will obtain for all whole number multiples (nf) of this frequency. At intermediate deflexion frequencies there will be a small residual effect on the beam, which will be a maximum at $(n + \frac{1}{2})f$, but this maximum decreases rapidly with the value of n . For frequencies less than the critical value f the sensitivity rises rapidly to a constant value at and below $f/2$; in other words, the deflecting frequency should not be higher than is sufficient to impress one half-cycle on the plates (or coils) during the passage of the beam through them. Tubes may be designed to operate at very high frequency by shortening the plate-length l and increasing the accelerating voltage, so that the time of passage of the beam is decreased. The deflexion sensitivity under such conditions is necessarily low (cf. VII.6).

Cathode-ray oscillographs for recording such high-speed effects are similar in principle to the low- and medium-voltage type discussed above, but very different in construction. They operate at 40–100 kV. and usually employ a simple cold cathode electron gun followed by magnetic focusing of the beam. The high speed of the electrons and the short deflecting plates allow the observation of wave-forms or transients in a range of

duration from 20 to $1/10$ microsecond. The angle of deviation of the beam, however, is correspondingly small, and it is necessary to have a large plate-film distance in order to get an image of readily observable size. Photographic recording is normally employed, and here again the high accelerating potential is of advantage, ensuring that enough electrons fall on the image to produce a photographic effect in the short tracing time. The film or plate is exposed directly to the beam in order to give higher recording sensitivity than would be possible if a fluorescent screen were interposed. The penetrating power of electrons at such speeds gives rise to high photographic efficiency, producing a visible image with much less beam current per sq. cm. than does a slower beam.

The introduction of photographic emulsions into the apparatus necessitates continuous evacuation, on account of their high vapour-pressure, even when the source of electrons is not a gas-discharge. To maintain steady production of the beam in this case a controlled leak V of air or inert gas is allowed into the discharge tube so that its pressure is of the order of 10^{-2} mm. Hg; it is usually connected with the rest of the apparatus only by a channel of very small bore through which the beam passes into the deflexion space (Fig. 116). With such low vacuum conditions prevailing, it is unnecessary to seek the same perfection of construction that causes a glass envelope to be employed for the normal sealed-off type of cathode-ray tube. It is now sufficient to use insulating material (glass or steatite) only for the discharge tube D and the bushes carrying connexions to the deflector plates. The main body of the instrument is made of metal, allowing of robustness of construction and safety in operation at very high voltage, the body being earthed and the cathode maintained at high negative potential.

A typical version of the high-speed oscillograph is shown in Fig. 116, the Metropolitan-Vickers type due to Burch and Whelpton. The cold cathode C delivers a beam of electrons through a fine hole in the anode A to a region in which concentration is effected by the magnetic focusing coil M . The trapping plates T, T' are normally biased with a potential sufficient to deflect the beam on to the plates P_x , so that the photographic plate at S is only exposed when desired. The pair of deflecting plates P_y carry the unknown impulse, and the upper plates P_x carry saw-tooth impulses of controlled frequency, giving a linear sweep (or 'time-base') to

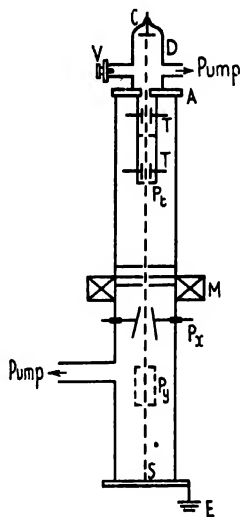


FIG. 116. High-speed cathode-ray oscillograph. (Burch and Whelpton.)

the beam in the X -direction. The preliminary adjustment of the beam is carried out under visual observation on a fluorescent screen S , which normally covers the photographic plate and can be raised when a record is to be made. The impulse under investigation is caused to lift the bias from the trapping system and thus release the beam, which is then deflected by the transient or high-frequency impulses applied to P , giving a corresponding trace on the emulsion. Evacuation of the apparatus is carried out by oil-diffusion pumps, or in other models by rotary molecular pumps, in order to exclude mercury entirely from the system. The Burch and Whelpton oscillograph operates at 15 kV. and is made in both single- and three-phase models. The Finch oscillograph is essentially similar in its construction, but operates at voltages up to 60 kV.; the general form is the same as that of the Finch diffraction apparatus (section 59, Fig. 119).

The electron optical arrangement of the usual cold cathode discharge tube is of the simplest, a rod of aluminium or magnesium-aluminium alloy serving as cathode and a plate of brass or copper as anode. In order to concentrate the emission into a small spot the cathode is sometimes made convex in shape or surrounded with a concave shield at the same potential. The returning beam of positive ions bombarding the cathode gradually results in the formation of a pit, around the edges of which emission is then concentrated. The cathode therefore has to be changed, or its surface repolished, after every few hundred hours of operation. A hot cathode was not favoured in the earlier instruments because of the rapid disintegration it suffered at the low vacuum which was achievable, even in the absence of the air-leak required by a discharge tube, and because of the difficulty of forming a beam of small initial cross-section without a complicated electrode assembly. The cold cathode also had the advantage of not requiring a filament heating circuit at high voltage. In any case, an oxide-coated cathode was out of the question as it would not withstand the electrostatic stresses set up at high accelerating potentials. More recently, however, progress has been made in designing a hot-cathode gun employing a tungsten wire bent to a V-shape, the point of which is the emitting surface, and a surrounding Wehnelt cylinder at like potential. A high pumping speed maintains a good vacuum and thus assists in prolonging the life of the filament, which is constructed so as to be easily removable. A beam is produced which is of readily variable intensity and steadier than that from a cold cathode, especially in the region of 100 kV.; the difficulties caused by the leak of gas from the discharge tube to the rest of the apparatus are automatically eliminated. At lower voltages the tungsten point may be coated with an oxide emitter. On the other hand, the simplicity of the cold cathode system has favoured its retention for most routine work.

FURTHER READING

1. The Cathode-Ray Tube:

Parr, *The Cathode Ray Tube and its Applications*, 1941.

Reyner, *Cathode Ray Oscillographs*, 1943.

Millman and Seely, *Electronics*, 1941.

de Gier, *Philips Techn. Rev.* 5, 245, 1940.

Pierce, *P.I.R.E.* 29, 28, 1941.

2. Television Cameras:

Campbell Swinton, *Nature*, 78, 151, 1908; 118, 590, 1926; *J. Röntgen Soc.* 8, 1, 1912.

Zworykin, *J.I.E.E.* 73, 437, 1933.

Lubszynski and Rodda, British Patent 442,666; 1934.

McGee and Lubszynski, *J.I.E.E.* 84, 468, 1939.

Iams and Rose, *P.I.R.E.* 27, 547, 1939.

Zworykin and Morton, *Television*, 1940.

3. High-speed Oscillographs:

MacGregor-Morris and Henley, *Cathode Ray Oscillography*, 1936.

Miller and Robinson, *Reports on Progress in Physics*, vol. II, 1936.

Wilson (W.) *The Cathode Ray Oscillograph in Industry*, 1943.

CHAPTER VIII

ELECTRON DIFFRACTION: THE ELECTRON MICROSCOPE

58. Electron Diffraction: Principles

IN discussing the behaviour of electron beams we have so far assumed the electron to be a corpuscle to which the ordinary methods of dynamics may be applied. Similarly, in the study of light the treatment of geometrical optics proceeds without reference to the ultimate nature of the rays considered. Any wave motion, however, will experience diffraction in passing through apertures of the same order of size as its wave-length. The electron exhibits wave properties and therefore may be diffracted. Its associated wave-length depends on its velocity and is less than an Angstrom unit when the accelerating voltage is 150 V. and upwards, so that diffraction is only possible in apertures of atomic dimensions. It was, in fact, the demonstration that electrons were diffracted by the regular array of atoms in a crystal lattice which first gave direct evidence of the wave nature of the electron.

As stated in section 50, the energy of a photon of radiation is given by

$$E = h\nu, \quad (\text{VIII.1})$$

where h is Planck's constant and ν is the frequency; that is to say, in regard to their energy light waves show corpuscular properties. In 1924 de Broglie put forward the hypothesis that conversely material particles might behave in some respects as waves. He suggested that their momentum, rather than their energy, was quantized and related to the equivalent wave-length (λ) by an equation analogous to VIII.1:

$$mv = h/\lambda, \quad (\text{VIII.2})$$

where m is the mass and v the velocity of the electron, and h again is Planck's constant. On inserting numerical values and expressing the motion of the electron in terms of its accelerating voltage V , we have for the wave-length (in Angstrom units)

$$\lambda = \sqrt{(150/V)}. \quad (\text{VIII.3})$$

Hence at the comparatively low voltage of 600 V. the wave-length is only $\frac{1}{2}$ Å, and at the practicable value of 60 kV. it is $1/20$ Å. Protons and more massive particles have a wave-length smaller in inverse proportion to their mass, and consequently diffraction of them is more difficult to observe. It must be noted that m , in equation VIII.2, is the instantaneous and not the rest mass of the electron. The relativistic correction (see p. 13) must be applied in accurate work when V is greater than 10 kV., giving an additional term in VIII.3.

Direct experimental verification of de Broglie's hypothesis by electron diffraction was provided independently by Davisson and Germer and by

G. P. Thomson. The diffraction effect may be discussed in terms of reflection from regularly spaced layers of atoms, as in the case of X-rays, or from rows of atoms. The latter corresponds more closely to the condition obtaining when electrons are transmitted through thin films, and will here be taken as the starting-point of discussion. In practice diffraction will occur from a three-dimensional grating, and not in a line grating as is usually the case with visible radiation.

A row of atoms, A_1, A_2, A_3, \dots , extends in the direction of travel of the electron beam E (Fig. 117). Scattering, or 'reflection', of the electron waves occurs from each atom, and these waves will arrive at a screen S ,

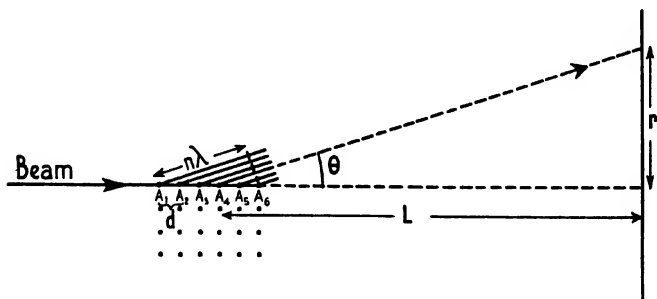


FIG. 117. Diffraction of electron beam at a row of atoms (atomic separation greatly exaggerated in comparison with L and r).

placed normal to the beam, with all possible phase relations. The condition for reinforcement, or maximum intensity, will be a path difference of $n\lambda$, where n is any integer; zero intensity should occur where there is a phase difference of π or a path difference of $(n + \frac{1}{2})\lambda$. If d be the interatomic distance, then the condition for reinforcement may be written

$$n\lambda = d \cos \theta, \quad (\text{VIII.4})$$

and will be satisfied at all points on the screen on a circle of radius $r = L \tan \theta$, if L is the distance of the screen from the point of diffraction. The diffracted beam of the first order leaves the crystal as a diverging cone of angle $2\theta_1$, and higher orders form similar cones of semi-angle θ_2, θ_3 , etc., as given by equation VIII.4. Bright rings are formed where these cones intersect the screen; the greater the number of atoms of spacing d which participate in forming the diffraction pattern, the brighter and sharper will be the rings.

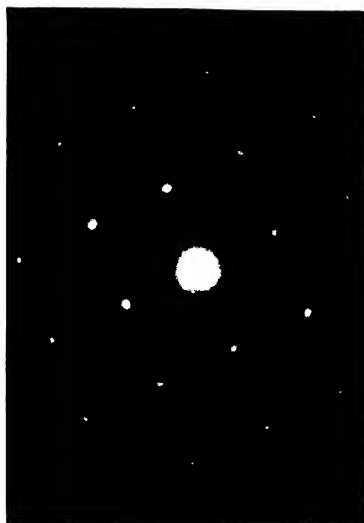
The crystal consists of a regular array of such rows of atoms in three dimensions, and their co-operative effect is found to give not circular rings but separate diffraction spots on the screen, in a pattern determined by the symmetry of the lattice (cf. Plates IV *a* and IV *b*). The addition of rows parallel to the first row, and in the same plane, will introduce a second periodicity d_2 along the direction normal to the electron beam. A

second set of diffraction cones will be produced at angles given by the analogous form of VIII.4, but with their axis normal to that of the set shown in Fig. 117. Diffraction maxima will therefore only appear on the screen where these cones intersect the rings due to the first spacing d : the result is a pattern of bright spots. The addition of atoms in planes parallel to that of Fig. 117, to form a three-dimensional grating, gives rise to a third set of diffraction cones about an axis perpendicular to the other two. Diffraction spots can now appear only where three cones intersect, a condition so rigid that a pattern is only obtained in special circumstances of alinement of the beam with respect to the lattice planes (Laue spots). Fortunately the specimens used in much electron diffraction work are of necessity so thin that they present effectively a two-dimensional array to the beam, and a 'cross-grating' pattern is readily obtained. If the lattice is of hexagonal structure, then three arrays of atoms intersecting at 60° are presented to the beam and six spots appear on the circumference of the diffraction ring of each order, but each row of atoms produces such a set and a regular hexagonal pattern is formed (Plate IV *a*). The form and size of the spots depend on the size of the beam and the perfection of the lattice. If the structure is cubic, four points are formed on each circle and the resulting pattern is square in form (Plate IV *b*). Structures of other symmetry produce more complicated patterns, the form of which gives direct evidence of the lattice structure.

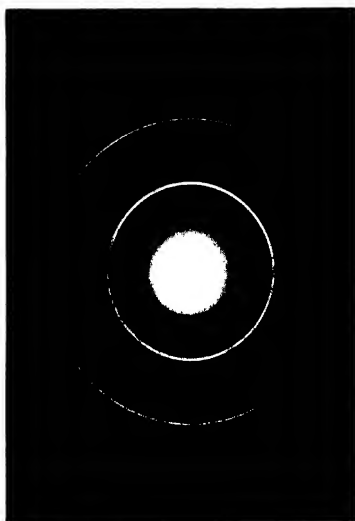
Such patterns can only be formed, however, if a perfect single crystal is responsible for the diffraction, and if it is orientated with its broad face normal to the beam. In general the specimen consists of a microcrystalline mass, and the pattern is more complicated. If the crystals are all parallel, but in random orientation about an axis in the beam, then the spots formed will lie around the circumference of circles having radii as given above, and in the extreme case will form one continuous ring for each order of diffraction (Plate IV *c*). More usually the arrangement will be completely random, so that different crystals will make different angles with the beam, and thus expose rows of atoms of different spacing to it. The values of d , determining the radius of the diffraction rings (VIII.4), will then take values corresponding to all the different sets of crystal planes present. A set of rings will then appear in the pattern, characteristic of the symmetry of the diffracting lattice (Plate IV *d*). If the crystal structure is known, the associated wave-length of the electrons may be calculated and equations VIII.2 and VIII.3 confirmed. Alternatively we may use electrons of known wave-length to investigate the crystal structure of chemical substances. All the methods employed in X-ray analysis are equally possible with electrons: single crystal rotation and polycrystalline powder observations being made. Surface reflection from solid objects as well as transmission through thin films yields an intelligible pattern. An essential practical limitation is the small penetrating power of the electron as compared with X-rays, the former being a particle with mass and therefore very readily absorbed and scattered by matter. Even



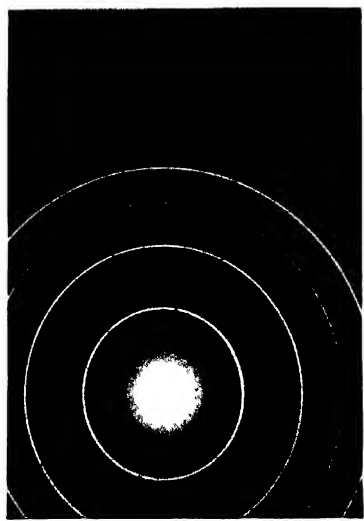
a. Cross-grating pattern of hexagonal structure (mica).



b. Cross-grating pattern of cubic structure (aluminum).



c. Ring pattern from chromium deposited on iron, showing partial orientation (strongly pronounced (111) rings).



d. Ring pattern from chromium deposited on bismuth, showing random disposition of crystals (very imperfect (111) orientation).

PLATE IV. ELECTRON DIFFRACTION PATTERNS.

(*Finch, Quarrell, and Wilman.*)

for beams accelerated by voltages of tens of kilovolts the specimen thickness may not be more than a few hundred Angstrom units without causing undue diffuse scattering of the electrons. Therefore, electron diffraction is primarily of value for the examination of thin films and surface layers, in which X-rays do not suffer sufficient diffraction to give a visible pattern. For the same reason the electron beam is diffracted by a jet of gas or vapour if the molecules have simple symmetry, as in carbon tetrachloride. In general, the efficiency of scattering of electrons is much greater than that of X-rays.

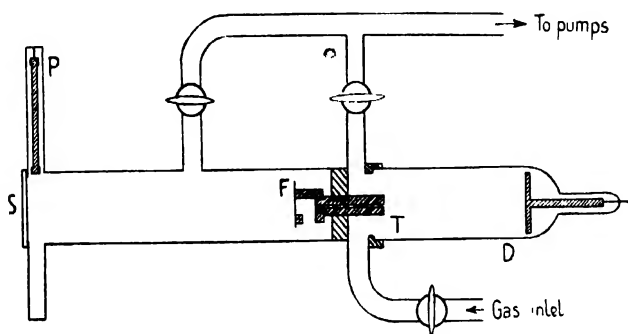


FIG. 118. Electron diffraction camera of G. P. Thomson.

59. *Electron Diffraction: Practice*

The original apparatus of G. P. Thomson is shown in Fig. 118. The electron beam was generated in the cold cathode discharge *D*, being limited by the channel *T* in its passage to the specimen *F* supported at right angles to it. The resulting diffraction pattern was observed on the fluorescent screen *S*, before which a photographic plate *P* could be exposed as required by lowering it from above. A slow leak of gas maintained steady conditions in the discharge, the apparatus being continuously evacuated. The source of high tension was an induction coil and the voltage was measured with a spark gap; up to 30 kV. were employed. Diffraction patterns were obtained from a number of metal films, and it was shown that the radius of the rings depended inversely on the root of the voltage applied to the electrons, as required by equations VIII.3 and VIII.4. Thomson then verified, by measurements on films of known structure, that the wave-length of the electrons agreed with that given by de Broglie's relation (VIII.2). Davisson and Germer produced similar evidence independently, from the diffraction of slow electrons from nickel single crystals. They used a Faraday cylinder as collector, in a fixed position, and observed the variation in current to it as the accelerating voltage of the beam was changed; the collector was then moved to a new position and the experiment repeated. Most subsequent work has used the Thomson method, with a constant high voltage and photographic

recording, the specimen usually being adjustable in one, and sometimes in three, dimensions.

Modern diffraction cameras make more use of electron optical technique.

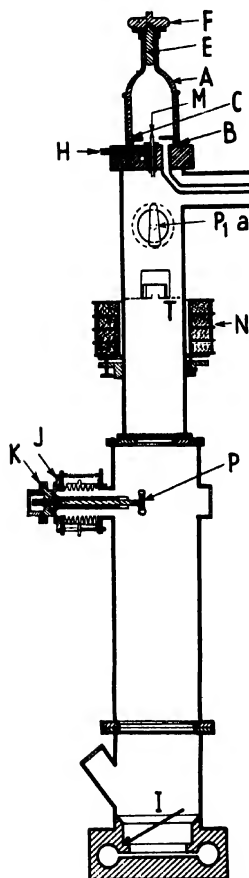


Fig. 119. Electron diffraction camera.
(Finch.)

The Finch (or Cambridge) apparatus is an outstanding example and will be described here (Fig. 119). It is mounted vertically, for convenience of operation, and is made on the lines of the high-speed cathode-ray oscillographs already described (section 57), being constructed entirely of metal apart from the discharge tube. The cold cathode *E* is surmounted by a corona ring *F* and is ground into the neck of the glass tube *A*. The anode *M* is a copper tube of fine bore, convex at the upper end to concentrate the beam, and is surrounded by a metal shield *B* for protection against X-rays. The discharge is maintained by a controlled leak *H*, evacuation taking place through the tube *L* by means of high-speed pumps. The anode block is water-cooled in order to allow high current densities to be used. The high tension (60 kV.) is supplied through a saturated diode, which ensures a constant current in the beam. A magnetic coil *N* focuses it to a fine spot on the observation screen *I*, which may be raised to expose a photographic plate or film. Whilst the discharge is being stabilized, the beam is deflected by a small potential applied to the plates *P*₁ and *P*₂, so that it falls in the beam trap *T*. The specimen is supported by the holder *P*, which may be rotated about three axes and also

translated by operating the ground joints *K* and screws *J* mounted on metal bellows. The diffraction pattern produced by the specimen in different orientations in the beam may thus be readily observed.

A number of workers have employed hot filaments for generating the electron beam, sometimes also using Wehnelt cylinders or other secondary electrodes for concentrating it. The use of a magnetic coil for final concentration of the beam is universal. Many types of specimen carrier and of recording chamber have been described. When it is desired to observe diffraction in gases, the former is replaced by a jet which delivers a

molecular beam across the electron path, to be condensed on a liquid-air cooled surface. Such special cameras have been described by Wierl, Cosslett, Brockway, and others. The regularity of atomic separation in a symmetrical molecule such as carbon tetrachloride gives rise to diffraction maxima and minima, but the random distribution of molecules in the gaseous state causes a diffuse background which severely limits the resolution obtainable.

The electron diffraction camera has thus become a research instrument of great importance for investigating the structure of matter, and particularly for the study of thin films, surfaces, and the molecular structure of gases. The electron optics of the apparatus is comparatively simple, and its operation a matter of no difficulty. The chief demands on skill arise in preparing specimens for observation and in interpreting the resulting diffraction patterns.

60. *Electron Microscopy*

The practicability of producing electron lenses of short focal length at once suggests their use in combination as a microscope. Indeed, the refinement of the magnetic lens has been due in large part to its development for electron microscopy; electrostatic lenses were developed primarily for cathode-ray tubes and have only recently been adapted to the severe requirements of short focus at high voltage. In exact analogy with light optical practice electrons may be used to illuminate (thin) objects and, in passing through a proper combination of lenses, to produce a greatly enlarged image. Magnifications of the order of 10^4 are readily obtained.

Resolving Power. Before describing the construction of such an instrument, however, it is necessary to make clear the limitations of light microscopy and the reason for the special advantages possessed by the electron microscope, which have rapidly made an important place for it in research laboratories. As is well known, the resolution of any piece of optical apparatus has a definite limit: two objects must subtend more than a certain minimum angle before they can be observed as separate points ('resolved') in the image. The limit is set by diffraction phenomena entirely similar to those discussed above. In terms of Fig. 120, each of two object points A , B will give rise to a diffraction pattern in the image plane, at A' , B' . These image points will only be observable as separate entities if the primary maxima of the two patterns do not overlap. The usual assumption is that the minimum permissible separation is one half-period: the central maximum of the one must fall at least at the first minimum of the other. For this to be so the paths to any image point from the two object points must differ by $\lambda/2$, or, by analogy with VIII.4,

$$\lambda/2 = \vartheta_0 \sin \theta, \quad (\text{VIII.5})$$

if θ is the semi-angle subtended by the object at the lens, the sine instead of the cosine now being involved, since the row of diffracting atoms is arranged normal to the beam, and not along it as in Fig. 117. The limiting linear

resolving power of the lens is then said to be equal to d_0 ; if A and B are closer than this distance, their image points will fuse and will no longer be separately distinguishable. When the refractive index of the object and image space is different, this becomes Abbe's relation:

$$d_0 = \lambda / (2N \sin \theta), \quad (\text{VIII.6})$$

where N is the relative refractive index of object and image space; $N \sin \theta$ is frequently written as NA , the numerical aperture of the lens. For optimum magnification it is usual in microscopy for the object to be situated very close to the focus of the lens, and we may write $D/2f$ for

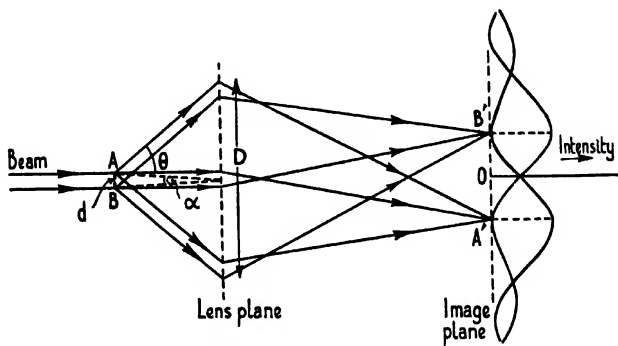


FIG. 120. Resolving power: limitation by diffraction effects. Diffraction at points A and B gives rise to diffraction rings about the image points A' , B' ; the intensity distribution is as shown when the images are just resolved.

$\sin \theta$, where D is the diameter of the lens and f its focal length. We then have for the angular resolving power (in radians)

$$\alpha = d_0/f = \lambda/D, \quad (\text{VIII.7})$$

a form which is convenient for use in respect of instruments of small resolution, such as telescopes and the eye.

The limit of resolving power thus depends on the numerical aperture, or diameter, of the lens and on the wave-length of the light employed. The normal human eye has a pupil of only a few mm. diameter and a correspondingly poor resolving power: it is of the order of 1 second of arc, or about 0.1 mm. at the distance of distinct vision. Finer detail than this cannot be appreciated by the unaided eye in objects or photographic images. In microscopy the best linear resolution will be obtained if the numerical aperture of the objective lens is high, that is, if $\sin \theta$ and the refractive index of the medium are both large. By using hemispherical condensing lenses practically the whole of the light transmitted by the object may be gathered, so that $\theta = 90^\circ$; and by immersing the object in a medium of the same refractive index as the condenser the value N is made uniform and approximately equal to 1.5. The resolving power may then be written

$$d_0 \approx \lambda/3.$$

The limit of resolution proves to be one-third the wave-length of the illuminating radiation, in the best type of light microscope. Even with the use of light from the blue end of the spectrum this limit will be of the order of 1,500 Å; with white light it is about 2,000 Å (2×10^{-5} cm.). Since the resolving threshold of the unaided eye is 10^{-3} cm., a magnification of 500 will then be sufficient to render visible detail of this order; greater magnification will result merely in enlargement of the ultimate image points without effecting further clarification. The highest power optical microscopes are therefore designed to give magnifications little greater than 1,000. By employing ultra-violet light, quartz lenses, and photographic recording it is possible to lower the limit slightly farther, but it is impracticable to use radiation of wave-length less than 2,000 Å. The maximum resolution so far achieved in an instrument of this type is 800 Å.

The Electron Microscope. A very large range of organisms, as well as nearly all chemical molecules, are of this order of size or smaller: viruses, many bacteria and bacteriophages, chromosomes and other units of cell structure. The light microscope makes visible a great range of biological and metallurgical structure, but fails at the borderline of the most important units of living matter. X-rays, being smaller in wave-length than the atom, permit investigation of the arrangement of atoms in molecular structures by true diffraction, but they cannot form direct images of a structure since they cannot be focused in lens systems. However, as pointed out in discussing electron diffraction, the electron has a wave-length of the same order as X-rays, when accelerated by more than a few hundred volts. Even if only small numerical apertures may be employed in electron lenses, it follows from VIII.6 that the resolution achievable should be substantially better than that of the best light microscope. In this expectation a magnetic electron microscope was constructed by Knoll and Ruska in 1931 and has been substantially improved since; its electrostatic counterpart has also been developed. Although nothing like the theoretical limit of 1 Å or less has yet been reached, resolving powers of 25–30 Å have been frequently claimed and a research instrument is now on the market with a guaranteed resolution of 100 Å or better.

The general lines of the electron microscope are shown in Fig. 121, side by side with a schematic view of a light microscope; the electronic type is normally arranged with the viewing screen at the lower end, and the light type has accordingly been inverted from its usual position, for convenience of comparison. The source of electrons *C*, usually a tungsten filament, provides a beam which is collimated by the condensing magnetic lens L_1 into a practically parallel beam, which falls on an object *O* supported immediately above the pole-pieces of the magnetic objective lens L_2 , of a few mm. focal length. The primary image formed by the latter is projected into the focal plane of a second strong lens L_3 (the projection lens) which forms a further magnified image on the fluorescent

screen S , beneath which is placed a photographic plate or film. The chief difference in form from the light microscope is in the comparative weakness of the condensing lens, due to the small lens apertures which may be used in the present state of correction, and in the use of a projection rather than an eye-lens, necessitated by the method of viewing. Otherwise the two systems are exactly analogous. The resolving limit being now of

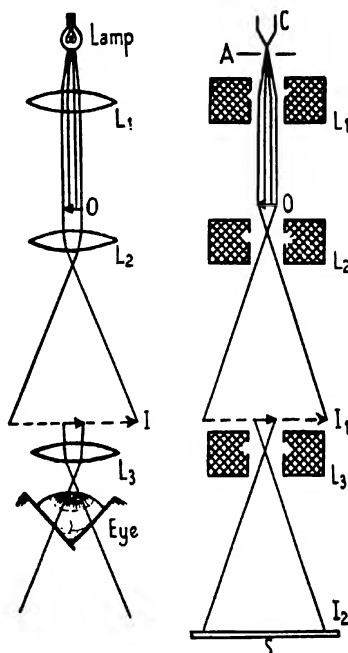


FIG. 121. Comparison of light and electron microscopes.

the order of 50 \AA , a total magnification of 2×10^4 will be required to enlarge such minute detail to the resolving threshold of the eye; this is achieved with a magnification of 100 in the first stage and 200 in the second stage. A smaller magnification may be used if recording is made on fine-grain photographic plates, the image on which is subsequently enlarged. It is undesirable to make the instrument very long, and if an image distance of 50 cm. in each stage is allowed, then the lenses must have a focal length of 2.5 mm. in order to give a magnification of 200. Such high-powered objectives are not difficult to design in the case of the magnetic lens, but it is found that spherical aberration is troublesome, with the result that very small apertures only are employed: of the order of 10^{-3} , as compared with the optical values of 1 to 1.5. All the lenses in the electron microscope of Fig. 121 may be replaced by electrostatic

lenses, and a number of instruments have been constructed of this type, symmetrical lenses being employed (section 62). Considerable difficulty arises, however, in making them of sufficiently small focal length at the high potentials required by other considerations, but the avoidance of corona discharge appears now to have been largely overcome, and there is a prospect that the advantages of the electrical lens may be exploited, in particular the small aberration of the univoltage lens.

The Specimen. The objects investigated require special preparation and mounting. They must be largely transparent to the electron beam, and therefore less than about 500 Å in thickness, except when accelerating voltages of 200–300 kV. are used. Sections of such fineness cannot be cut by the normal type of microtome, although Ardenne has devised a type which cuts sections of wedge shape with such precision that they fall below this limit at the thinner end. Otherwise the objects must be suspended in water or other liquid and carefully deposited on extremely thin films of a cellulose nature; after drying, they are mounted in a special holder and inserted into the vacuum of the microscope. The necessity of drying specimens and exposing them to a vacuum prevents at present the investigation of living material. On the other hand, it is possible by a special casting technique to observe the structure of solid surfaces. Direct reflection of electrons from, say, a metal surface does not yield an intelligible image. The method employed is to deposit on it a thin layer of a structureless substance such as a resin or plastic, which may then be stripped off or released by dissolving the metal in acid. A few metals, such as nickel and aluminium, form oxide coatings which can be stripped from the surface without destroying their microstructure. The best results have been obtained with a casting method in which polystyrene is first pressed on to a metal surface at 160° C. and under about 100 atmospheres pressure. The metal specimen is then sawn off short and the rest dissolved with acid, the polystyrene mould is mounted *in vacuo* and a thin layer of quartz very rapidly deposited on it by evaporation. After dissolving away the plastic in ethyl bromide, the silica replica remains and is mounted for examination in the electron microscope. It gives clearer definition in the ultimate image than does the original polystyrene mould itself. Heidenreich and Peck obtained a resolution of about 100 Å from replicas of metal surfaces made by this two-stage process; an example is shown in Plate VII a.

Image Formation. A word is necessary concerning the details of image formation in the electron microscope, since they differ from those obtaining in the light microscope in a manner that materially affects the problem of lens design. In a light beam the structure of a specimen is made visible primarily by differential absorption of the radiation in different regions, aided frequently by the application of special staining materials. In the electron microscope, however, the absorption of electrons (in the strict sense of the term) may vary little in the different parts of a specimen. The main cause of diminution of intensity of the beam lies rather in

differential scattering from molecules of varying structure and chemical nature. Regions transparent to electrons cause negligible change in the local beam intensity, on a sufficiently thin carrier. Materially denser regions scatter more or less electrons from the incident beam, with the

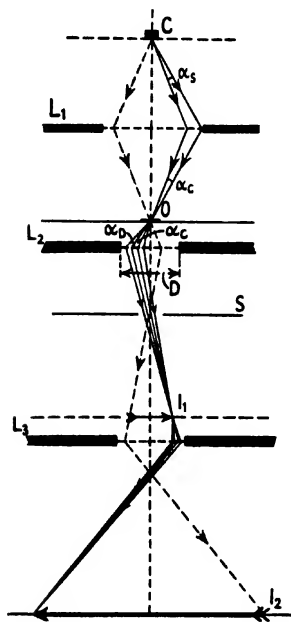


FIG. 122. Image formation in the electron microscope.

result that corresponding areas in the image are less brilliant or may be entirely unilluminated. The scattered electrons form a diffuse background to the image, or may be lost from the final beam altogether, since the magnification is very great. It follows also that only a portion of the field in the objective lens, and a still smaller portion in the projection lens, is concerned in the imaging of a given object point, the effective aperture thus being less than that given by simple geometrical considerations and the aberrations of the lens being to that extent less effective.

The path of the bundle of rays concerned in the imaging of a single object point is shown in Fig. 122. The condensing lens collects a beam of aperture α_s from the electron gun and forms it into a convergent illuminating beam of aperture α_c determined by the position of L_1 between gun and object. This beam is scattered by an amount dependent on the mass-thickness of the object O , and

also diffracted at its edges. The position of the image point in the primary image I_1 is determined by the point of intersection of the unscattered rays of emergent aperture α_c , as shown. Those electrons which are scattered elastically will be imaged in a circle of confusion around this image spot, the size of which will depend on the degree of aberration displayed by the lens. To reduce this effect a very small aperture stop S is introduced at the plane of minimum cross-section of the emergent beam, corresponding to the cross-over of an electron gun, as indicated. According to the Abbe diffraction condition, at least the full width of the first maxima must be focused if a clear image of the edges of the object is to be obtained. The stop must therefore have an aperture not less than that of the incident beam, α_c , plus twice the aperture of the diffracted pencil, α_D . The latter is given by VIII.5, on insertion of the wave-length of the illuminating electrons and the diameter d of the object point. The width of the aperture stop must thus at least correspond to $\alpha_c + 2\alpha_D$. If the brightness of the image is sufficiently large, and the value of d is as small as desired, there is no virtue in using a larger aperture, since the result

would be an increase in the confusion of the image owing to spherical aberration. The use of a small stop also eliminates most of the inelastically scattered electrons, that is, those which lose energy on collision. These slower electrons are deviated more strongly in the lens than those of unaltered speed, and will thus not fall into the aperture. Such as reach it from other object points will mostly be travelling on paths which fall outside the aperture of the projection lens L_3 , and will be eliminated at this stage; the early microscopes, which did not employ an aperture stop, probably owed their good definition largely to this chromatic effect. A proportion of scattered electrons will, however, reach the final image I_2 both from this cause and from scattering at the rim of the aperture or lens diaphragms, but they will be distributed over the image and form only a diffuse background. The small degree of confusion is seen from the illustrations reproduced.

Apart from spherical and chromatic aberration, lens errors play a comparatively small role, as mentioned in the earlier discussion of aberrations. The spherical error predominates, although the chromatic effect is by no means negligible, especially at low accelerating potentials (see p. 140). Together they set an upper limit to the angular aperture which may be employed if a given resolution is to be obtained in the image, whereas the diffraction effect sets a lower limit. After a full investigation of the matter Ardenne lays down the rule that for optimum resolution the size of the aperture should be such that the total aberration effect is roughly equal to the diffraction error. The total aberration in the image plane and its variation with aperture are given by equations V.20 and V.34, and division by the magnification M gives at once the corresponding lack of resolution in the object. The variation in the angle of the diffracted cone with object size is given by the Abbe rule, VIII.6. Comparison of the two curves gives the aperture for which the object size equals the limit set by aberration. In this way is found the optimum angle (α_c) of the illuminating beam for any object point, and thus the operating conditions of the condensing system. The total extent of the diaphragm in the objective lens L_2 will then be governed by the size of specimen to be imaged, and this in turn is determined by the maximum permissible size of the final image. In practice a diaphragm opening D of about 0.1 mm. is employed in a lens of focal length 2 mm., the effective aperture being of the order of 10^{-2} mm. in diameter and thus the angular aperture approximately 5.10^{-3} .

It will be clear from Fig. 122 that similarly only a very small portion of the field of the projection lens L_3 takes part in the focusing of an individual pencil. The aberration of an image point is thus much less than would be concluded from measurement of the distribution of the field, with the light optical analogy in mind. Entirely the same considerations apply to the imaging action of electrostatic lenses, the main practical difference being that the chromatic error is much less when they are of the univoltage type.

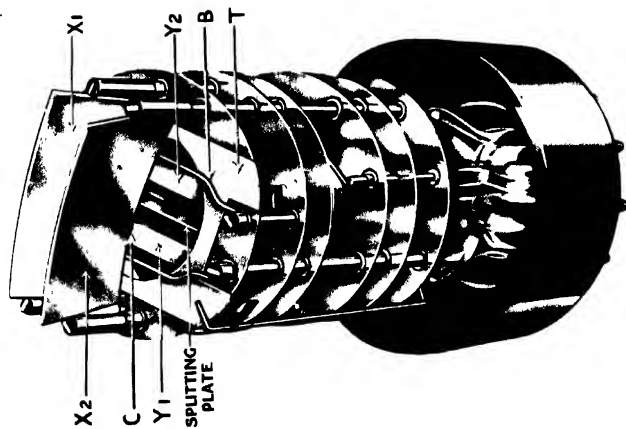
61. *The Magnetic Electron Microscope*

The first electron microscope was described by Knoll and Ruska in 1932, and was of the magnetic type, being similar in its essentials to that shown in Fig. 121. Subsequent independent developments by Marton in Liège, and by Hillier and Prebus in Toronto, led to a much improved model being produced when these workers joined Zworykin in the R.C.A. laboratories. In Great Britain Martin's early instrument has now been refined by Metropolitan-Vickers and will be described here.

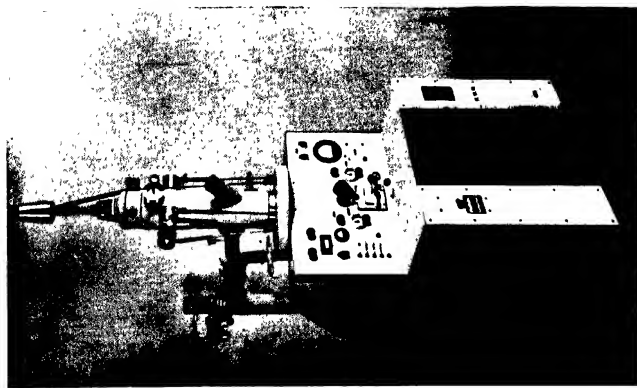
A section through the microscope is seen in Fig. 123, where E is the electron gun employing a tungsten point as hot cathode and a grid shield G at the same potential. The lenses L_1 , L_2 , and L_3 are so designed that they carry adequate energizing current without need of cooling. A cleverly contrived object chamber O enables a specimen, in a special cartridge, to be introduced into the correct position above the poles of the objective lens without more than a slight disturbance of the vacuum conditions. High-speed oil-diffusion pumps maintain a pressure of the order of 10^{-5} mm. throughout the apparatus, large ports P_1 allowing a high pumping speed. The photographic plate or film is located at B in the base of the instrument; it can be racked along by external control in order that several successive exposures may be made. The bevelled surface of the lower chamber carries pairs of viewing ports V which allow binocular vision of the final image, seen on the fluorescent screen which normally covers the plate. An intermediate viewing screen S_1 is available for observation of the first image when adjusting the specimen in the beam. The total electronic magnification is 10,000 times, and provision is made for a further optical or photographic enlargement of 10 times. The main body is made of metal in several sectional units, for ease of assembly and demounting, and is earthed. The whole of the path of the electron beam must be completely shielded against stray magnetic fields, and this is effected by tubes of Mumetal or some other material of high permeability. The high-tension supply, operating at 50 kV. or lower, is contained in a separate unit from which leads run to the electron gun and to the windings of the lens coils. Meters recording the beam current and accelerating potential, the coil currents, and the vacuum pressure are conveniently mounted on a desk panel together with the controls. The whole system is comparatively compact and easy to handle, so that it can be considered a finished research tool.

The R.C.A. high-resolution model is similar in its essentials, except that the high-tension unit and controls are contained in the same assembly as the microscope; a console model is also made (see below, and Plate VI a). An electron diffraction adapter is provided, which permits micrograms and diffraction patterns to be obtained from the same specimen. It takes the place of the normal projection lens L_3 , and comprises a projection lens, specimen holder, and focusing lens.

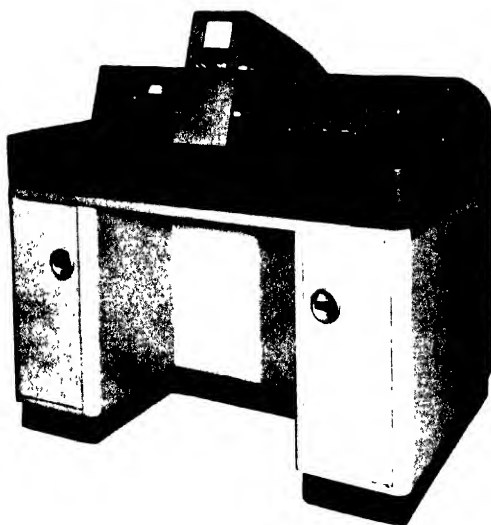
Both instruments give a resolving power of 100 Å, or better, in their



a. Plate structure of double-beam cathode-ray tube. The Y-plates are separated by a splitting plate at earth potential. The curved edge *C* of the X-plates and the shield *T* correct distortion; the wires *B* reduce inter-modulation ('cross-talk').
(A. C. Cossor Ltd., London.)



b. Electrostatic electron microscope.
(Dr. P. Grivet, Paris.)



a. Console model magnetic electron microscope.
(Radio Corporation of America.)



b. Magnesium oxide smoke ($\times 8,000$).
(Dr. P. Grivet.)



c. Etched aluminium surface (oxide replica)
($\times 8,000$). (Dr. P. Grivet.)

commercial form; the prototypes in experienced hands have given results down to 30 Å. Typical photographs of bacteria and oxide smokes are shown in Plates VI and VIII, which include reference scales to indicate

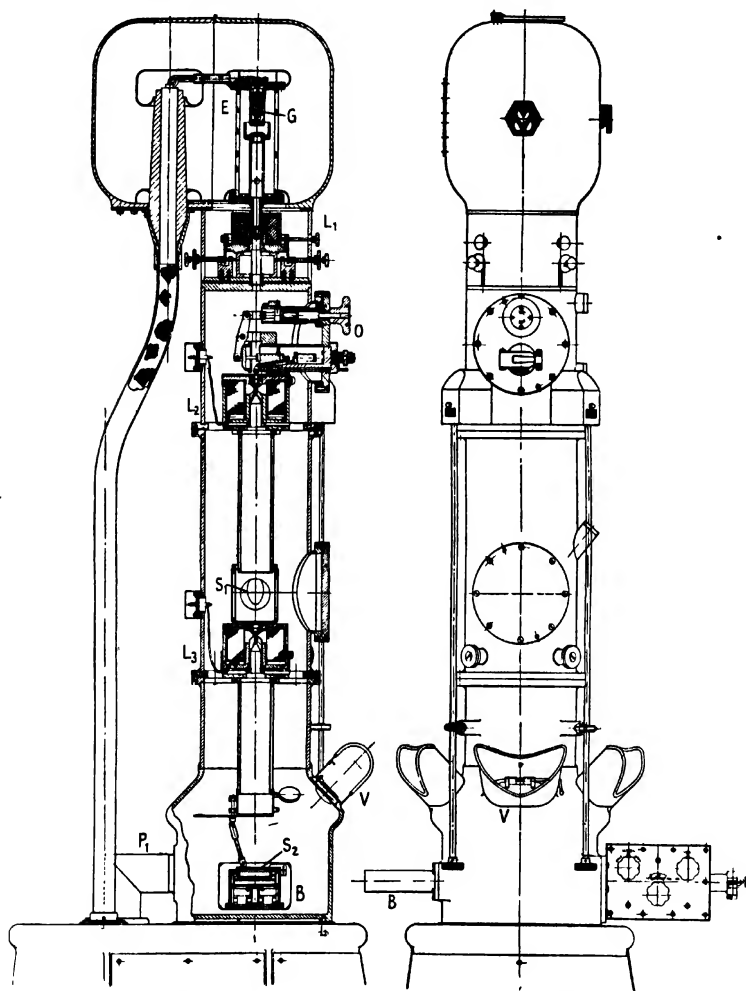


FIG. 123. Magnetic electron microscope. (Metropolitan-Vickers.)

the degree of resolution. A magnetic microscope of similar efficiency has been described by Ardenne, using an accelerating voltage of 200 kV. in order to obtain greater penetration of the specimen; R.C.A. have operated an instrument at 300 kV.

For many purposes in a research laboratory a lower magnification than 10,000 is sufficient, covering the order of resolution immediately

beyond the limit of the light microscope: from 200 to 2,000 Å. To meet this need the R.C.A. have produced a smaller magnetic instrument, and the G.E.C. an even more compact electrostatic microscope. The R.C.A. instrument is mounted nearly horizontally on a desk, for convenience of operation, and is shown in section in Fig. 124. In order to reduce the total length the condenser lens is eliminated and the other two lenses are

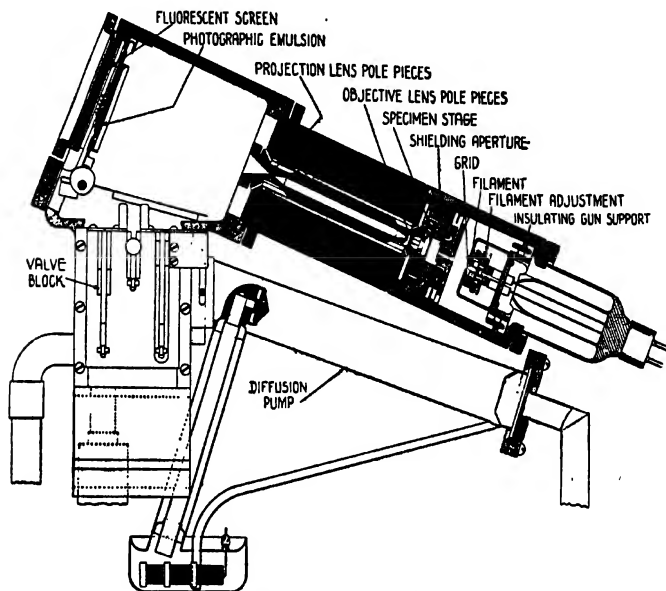


FIG. 124. Desk model magnetic electron microscope (vertical section). (R.C.A.)

combined in one unit. Adequate collimation of the beam is obtained by using the filament shield as a control grid, biased at a potential ranging from 0 to -75 V. with respect to the filament. The objective and projection lenses, L_1 and L_2 , are energized by a common winding encased in an iron shroud (heavy lines) except for the gaps between the two sets of pole-pieces at either end. A removable inner pole-piece is fitted to the projection lens, which has the effect of intensifying the focusing field. In its absence the total magnification is 500, but becomes 5,000 when it is inserted; the magnification otherwise is fixed in this model. The final image is received on the fluorescent screen S , outside which a large glass lens may be used for enlarged vision of the image. A photographic plate P may be exposed before the screen as required. The microscope is operated at 30 kV., and at the larger magnification gives a resolving power of 100 Å at maximum. The pumping unit and other apparatus is contained in the same desk, the controls being reduced to a minimum for ease of operation.

The main features of short-focus magnetic lenses have already been dis-



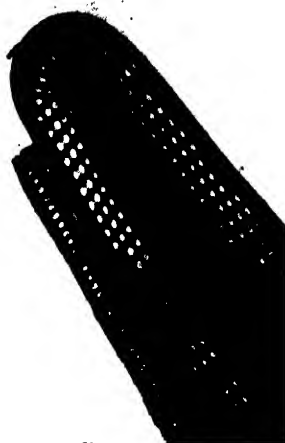
a. Cleavage face in fractured magnesium crystal (silica replica) ($\times 6,500$). (Dr. R. D. Heidenreich.)



b. Malarial parasite (*Plasmodium cynomolgi*) in a red corpuscle of monkey blood ($\times 15,000$). (J. Smiles: Nat. Inst. Med. Res., London.)



a. *Aucuba mosaic virus*
($\times 14,000$). (National Physical
Laboratory, London.)



b. Diatom: *Amphipleura pellucida*
($\times 10,000$). (National Physical
Laboratory, London.)



c. Zinc oxide smoke ($\times 20,000$). (Metropo-
litan-Vickers Ltd., Manchester.)



d. Bacteria (*Rickettsia*) of epi-
demic typhus (Cairo strain)
($\times 15,000$). (J. Smiles: Nat.
Inst. Med. Res., London.)

cussed in section 33 and in the treatment of aberrations (especially in sections 39 and 43). Spherical aberration is kept to small proportions by using very small apertures and high accelerating voltages (cf. equation V.23), both circumstances also contributing to the reduction of chromatic aberration. But, as explained in section 43, it is inherently impossible to correct a lens for the latter defect, and therefore it becomes essential to

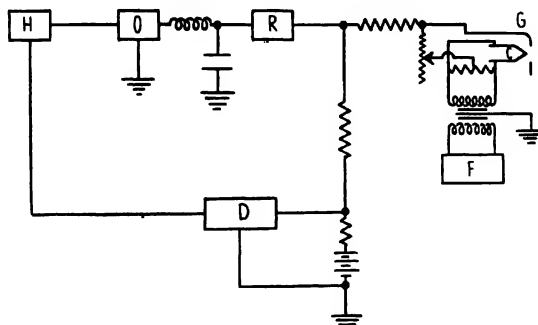


FIG. 125. High-voltage rectifying circuit for electron microscope. (Hillier and Vance.)

reduce the fluctuation in applied voltage to very small proportions. If the chromatic error is to be less than the spherical error, affecting the resolution by an amount of the order of tens of Angstrom units only, the voltage must be stabilized to an accuracy of 1 in 20,000. This is probably the most difficult technical problem in the construction of an electron microscope of the magnetic type. It was solved by Hillier and Vance by the application of modern radio technique to high-tension supply. The outline of the resulting system is shown diagrammatically in Fig. 125. A high-frequency voltage is generated in the self-excited driving oscillator *O*, which is regulated by negative feed-back. The oscillator output, of frequency 32 kc.p.s., passes from the series-resonant circuit of high step-up ratio ('*Q*-factor'), to the doubling rectifier *R* which delivers 30 kV. direct current to the microscope electron gun; at the high frequencies employed the output can be smoothed with high efficiency, especially as the high selectivity of the resonant circuit admits a very pure wave-form to the rectifier. Regulation is achieved by passing a sample voltage, from a voltage divider in the rectifier, through a two-stage D.C. amplifier *D* back to the regulator unit *H* which controls the high-tension anode input to the driving oscillator. The resulting stability of voltage is in the region demanded, 1 part in 25,000. Similar constancy of supply must obtain for the current through the winding of the magnetic lenses. It is pointless to seek much greater accuracy, since the variation in the velocity of thermal emission of electrons from the cathode is of the same order. In comparison with spherical aberration, the chromatic error is thus largely eliminated; it will demand further attention, however, as soon as means are found of reducing the former error.

62. *The Electrostatic Electron Microscope*

The use of electrostatic rather than magnetic lenses in the electron microscope offers certain definite advantages. No limitation is placed on the field strength by saturation effects, and also the chromatic error may

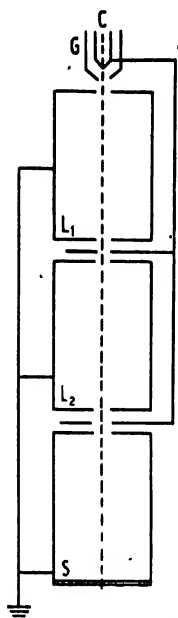


FIG. 126. Electrostatic electron microscope (schematic).

be reduced to negligible proportions if the uni-voltage type of lens is employed. The general form of such a system is shown in Fig. 126, the central elements of each of the two lenses L_1 and L_2 being at the same potential as the tungsten cathode C of the electron gun; the outer, symmetrical elements of the lenses and the body of the instrument are earthed. With respect to the electron, the second potential V_2 in each lens is therefore always zero and the focal length fixed. Any increase in velocity of the electron caused by a rise in cathode voltage is counterbalanced automatically by an increase in the power of each lens. Chromatic aberration due to fluctuations in the supply voltage (apart from a relativistic effect discussed below) is thereby eliminated, leaving only the small inevitable variation in velocity of thermal emission.

On the other hand, the technical difficulties of constructing electrostatic lenses of very short focus are greater than for magnetic lenses. The main difficulty lies in designing the shape of apertures so that high potentials may be safely applied to them when their separation is of the order of millimetres only, as is necessary if a focal length of this order is to be obtained. The problem was first solved by Mahl, and independently by Ardenne. Another solution was recently evolved by Bachman and Ramo, and forms the basis of the compact microscope already mentioned. The form of lens electrodes used by Mahl and by the latter workers are shown in Figs. 127 *a* and *b* respectively.

The microscope of Bachman and Ramo employs three stages of magnification, with the same type of lens in each. No condensing lens is used, the optimum aperture of the illuminating beam being fixed by careful adjustment of the filament in its shield and of the position of two limiting apertures. The hot cathode gun is shown in Fig. 128, the shield G being at the same potential (35 kV.) as the cathode, AA being the apertures at anode potential (earth), and MM being magnetic shields to protect the focusing system from the fields set up by the filament heating supply; the object is situated at O . The three symmetrical lenses are arranged in a single cartridge, the end of which carries the object stage and adjusting mechanism. The focus being fixed, focusing must be carried out by axial

movement of the object. The final image is projected on to a small fluorescent screen of the finest possible grain, where it is viewed or photographed with the aid of a light microscope. This combination of electronic and optical magnification was chosen in order to facilitate the construction of

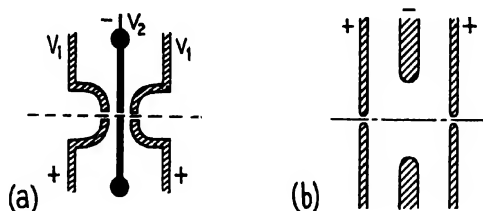


FIG. 127. Short-focus electrostatic lens. (a) Mahl; (b) Bachman and Ramo.

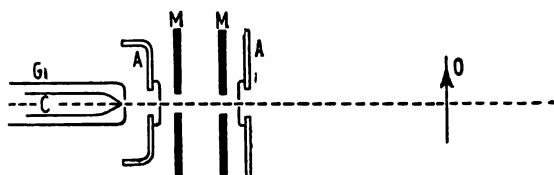


FIG. 128. Electron gun of electrostatic electron microscope. (G.E.C.)

the lenses and to reduce the size of the instrument. The total length from cathode to screen is no more than 11 inches. The overall magnification, including a sevenfold optical enlargement, may be varied from 3,000 to 8,000 by changing the spacing of the lenses; the highest resolution obtained is of the order of 200 Å.

The statement that microscopes of this type are entirely independent of fluctuations in the voltage supply is valid only to a second approximation. Ramberg has pointed out that a disturbing effect must be taken into account when the electron velocity is such that the relativistic change in mass becomes appreciable. An increase in voltage, producing an increase in v and therefore also in m , then results in a slight increase in the focal length of the symmetrical lens. Under the conditions of operation of existing models, however, this effect only sets the wide limit of 1,000 V. to the permissible voltage fluctuation; fortunately also the extent of the aberration decreases with increasing focal power for lenses having a negative central electrode.

The higher voltage electrostatic microscopes constructed by Mahl and by Ardenne have given resolving powers down to 60–80 Å, about twice that of the magnetic type. So far, less attention has been paid to the development of the former as compared with the latter type, and very possibly it may eventually provide the higher resolution. In both cases further advance depends chiefly on the improved correction of lens errors, and in particular of spherical aberration.

63. Other Types of Microscope

The field emission, or point projector, electron microscope has already been described (section 49). It involves no electron optical system, other than the provision of a high rectilinear field at the cathode, and does not belong in the present discussion. A somewhat different

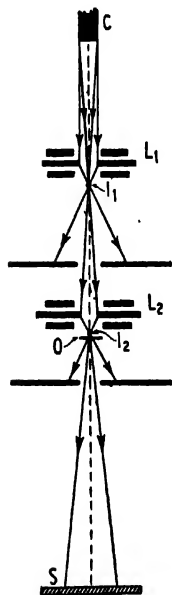


FIG. 129. Electrostatic projection microscope. (Boersch.)

type of projection microscope has been described by Boersch, which employs a fine beam of electrons to cast a shadow of an object on to a viewing screen; the fine beam corresponds to the point source of electrons in the field emission type, but is produced by demagnification of a larger beam in a number of electron lenses. The system is shown schematically in Fig. 129 where C is the cathode from which an electron beam is projected into the first electrostatic lens L_1 . The beam is brought to a focus close to the lens so that a much demagnified image is produced, which in turn acts as object for the second lens L_2 . The final image point I_2 is of the order of 100 Å in diameter; owing to the high degree of demagnification ($1:10^4$ or $1:10^5$) the beam current is now reduced to the order of 10^{-13} amp. I_2 acts as a point source of illumination for the specimen O , throwing a shadow-graph of it on the fluorescent screen S . Magnification is thus produced by simple projection, its magnitude being given by the geometry of the arrangement, as in the field emission type; the essential difference is that any thin specimen may now be imaged, whereas the latter microscope can only project the cathode surface itself. By careful adjustment of the position of the

specimen magnifications greater than 10^4 may be obtained, but the resolution is limited by the size of the image point I_2 . Attempts to reduce its diameter below 100 Å lead to such a reduction in beam current that the recording time becomes unduly long. Here again progress waits upon better correction of the lenses, so that beams of wide angle may be focused. Resolving powers of the order of 400 Å have been reported with present technique.

Ardenne has utilized a demagnified beam for another method of microscopy, in which the fine point source (or 'electron probe') is scanned across the object and the transmitted intensity is recorded photographically or electrically. The same type of deflexion system as in the cathode-ray tube is used for the scanning process, but at a much lower frequency. The time of scan is about 1 sec. per line and 20–30 min. per picture, owing to the low beam intensity. The resolution depends on the fineness of the probe and on the number of lines in the scan; Ardenne reports values down to 400 Å. The method presents the possibility of investigating living

tissues in air, since the specimen may be placed on the outer surface of a thin metallic window which is scanned on the inner side by the electron probe. Owing to the degree of scattering which occurs, however, high resolution in the recording of the transmitted intensity is difficult to obtain.

The scanning principle has been employed in the R.C.A. laboratories in a different manner: for the direct investigation of surface structure, which is otherwise inaccessible to the electron microscope, except where the plastic replica technique can be applied. A fine electron probe I_2 is produced by two stages of demagnification with electrostatic lenses (L_1 and L_2 , Fig. 130). A third lens projects an image of this source, at unit magnification, on to the surface of the specimen O , across which the beam is scanned by the deflecting influence of magnetic fields M . Electrons are diffusely reflected from the surface to an extent determined by its structure, and these secondary electrons are collected on the fluorescent screen S , the illumination from which produces a photo-electric response in the photocathode P of an electron multiplier. The input, and therefore also the output, of the latter is thus proportional to the local intensity of reflection of electrons from the surface at each instant of the scan. After further amplification this current modulates the synchronously-moving printing bar of a facsimile recorder, which thus produces a representation of the surface structure. Such a complicated system requires greater care in operation than the electron microscope and so far has not proved capable of comparable performance. Zworykin, Hillier, and Snyder have obtained reproductions showing a resolution of 500 Å, a considerable improvement over the best optical microscope but not so refined as the results of the replica method of surface investigation.

It remains to mention a further possibility of observing living specimens in air, rather than *in vacuo*: to use a stationary electron probe as a point source of X-rays, which then cast a shadow of the object on to a screen or plate. The differential scattering of X-rays by the structure of the specimen gives rise to a true image by geometrical projection, against a diffuse background. The object is again fixed across a thin window sealing the vacuum in which the electron beam is produced. Comparatively low accelerating voltages are necessary ($\sim 10,000$ V.) in order to avoid a broadening of the point source of X-rays in the window, and so that the

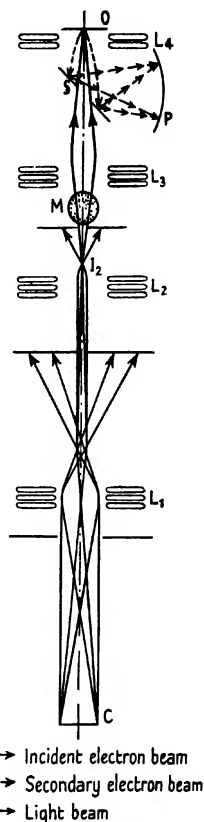


FIG. 130. Electrostatic scanning microscope. (R.C.A.)

rays themselves shall be of medium wave-length and comparatively easily absorbed. Such an arrangement is described by Ardenne, but it is not clear what results, if any, he has obtained with it. The resolving power is not likely to be of the same order as that of the electron microscope, but it offers a possibility of observing living organisms which lie immediately outside the range of the best light microscopes.

FURTHER READING

1. Electron Diffraction:

Thomson (G. P.) and Cochrane, *Theory and Practice of Electron Diffraction* (1939).

Symposium, *Trans. Far. Soc.* **31**, 1043–1290, 1935.

G. I. Finch, *Reports on Progress in Physics*, **2**, 291, 1935 (cameras).

2. Electron Microscopy:

H. J. H. Starks, *Reports on Progress in Physics*, **2**, 283, 1935.

O. Klemperer, *ibid.* **7**, 107, 1940.

Burton and Kohl, *The Electron Microscope* (1942) (elementary).

Ardenne, *Elektronen-Übermikroskopie* (1940).

Marton and Sass, *J. Appl. Phys.* **14**, 522, 1943; **15**, 575, 1944; **16**, 373, 1945 (complete bibliography).

Heidenreich and Peck, *J. Appl. Phys.* **14**, 23, 1943.

3. Magnetic Microscope:

Knoll and Ruska, *Z. techn. Phys.* **12**, 389, 1931; *Z. Phys.* **78**, 318, 1932.

Marton, *Nature*, **133**, 911, 1934; *Phys. Rev.* **46**, 527, 1934.

Martin, *Nature*, **142**, 1062, 1938; *J. R. Micr. Soc.* **59**, 217, 1939.

Prebus and Hillier, *Canad. J. Res.* **17**, 49, 1939.

Hillier and Vance, *P.I.R.E.* **29**, 167, 1941.

Zworykin and Hillier, *J. Appl. Phys.* **14**, 658, 1943.

Cossalett, *J. Sci. Inst.* **22**, 170, 1945.

4. Electrostatic Microscope:

Mahl, *Z. techn. Phys.*, **20**, 316, 1939; **21**, 17, 1940.

Bachman and Ramo, *J. Appl. Phys.* **14**, 8, 69, and 155, 1943.

5. Other Types:

Boersch, *Z. techn. Phys.* **20**, 346, 1939.

Ardenne, *loc. cit.*

Zworykin, Hillier, and Snyder: *Bull. Amer. Soc. Test. Mater.*, Aug. 1942.

CHAPTER IX

OTHER APPLICATIONS: CYLINDRICAL FIELDS

64. *Beta-ray Spectrometry*

AN important branch of the study of radioactivity is the investigation of the beta-rays, or fast electrons, which are emitted in the process of atomic disintegration. They exhibit a considerable variation of intensity with velocity both from a given atom, and from one atom to another. It is desirable to have detailed information of this velocity distribution, and particularly of its upper limit, for each stage in disintegration. Just as in the emission of visible light, it is found that sharp maxima occur at certain velocities, and this analogy gives the subject its name of beta-ray spectrometry.

Semicircular Method. The simplest form of beta-ray spectrometer utilizes the deflecting properties of a uniform magnetic field. If a line source of electrons, S , is set up along the direction of the field (Fig. 131), the emitted electrons will describe circular paths in a plane normal to this direction and more complicated paths in any other plane (cf. section 3). The radius ρ of the trajectory will be given by equation I.1:

$$H\rho = mv/e, \quad (\text{IX.1})$$

if H is the field strength, v the velocity, and e the charge of the electron; m is again its momentary mass, and the relativistic correction must be applied at velocities greater than 10^{10} cm. sec.⁻¹ All electrons emitted normally from the line source with the same velocity v will describe semicircular paths and again intersect the plane containing the source in some line I , which may be regarded as the image of the source S ; but no true focusing obtains in the sense of the concentration of a bundle of rays emitted within a given angle. The velocity distribution of the electrons may at once be found by varying the strength of the deflecting field whilst recording the number received at any fixed position such as I . The means of recording may be photographic, when the intensity of emission is high enough to give a visible image in a sufficiently short time, or may be an ionization chamber or Geiger-Müller counter when the emission is small. The source-image distance SI ($= 2\rho$) is such that a uniform magnetic field can be established across it without undue difficulty, whilst not so small that direct passage of radioactive radiations from S to I may not be prevented by interposition of lead blocks: it is usually of the order of centimetres rather than decimetres. Owing to the ease of absorption and scattering of electrons, the whole path must be *in vacuo*.

The line source will in practice have a finite width, and the image formed by electrons of the required velocity v_z will have an equal width if the field is perfectly homogeneous; a number of semicircular paths, having centres

in close proximity, now lead from the one to the other. In addition, however, a proportion of beta-rays of other velocities will be received in a recording slit of this given width, after travelling paths such as those shown in Fig. 131. In order to limit the beam, and thus obtain a high resolution in the recording, diaphragms must be sited along the semicircular path. For this reason the source is usually placed at S' below the diametrical

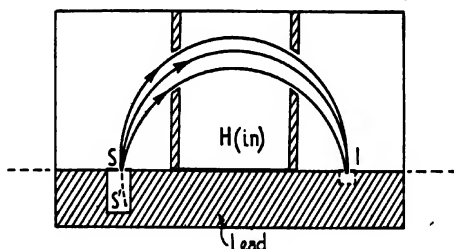


FIG. 131. Semicircular method of focusing beta-rays (magnetic field H directed into diagram).

plane and the line S becomes a limiting slit. Simple geometrical considerations then allow the resolution dv/v ($= d(H\rho)/H\rho$) to be calculated; it is usually of the order of 1 per cent.

The semicircular method has been the standard means of beta-ray analysis since the early days of radioactive investigations; it will be observed that the system has (semi-)cylindrical symmetry. A numerical evaluation of equation IX.1 shows that with a radius of 5 cm. a field strength of only 120 gauss is required for focusing electrons of velocity 10^{10} cm. sec.⁻¹. At a velocity equal to 0.9 that of light the equivalent ($H\rho$) value is 3,500 gauss-cm., and the required field strength is still not difficult to obtain. Beta-rays of this velocity correspond to electrons accelerated by a potential difference of 660 kV. Most beta-ray emission from naturally radioactive substances occurs with lower velocity; a typical spectrum is shown in Fig. 132, obtained from thorium-B. A number of sharp 'lines' appear in the emission, superimposed upon a continuous background which stretches from zero up to a sharply defined maximum velocity. Theoretical interest attaches primarily to the position and intensity of the lines and to the upper limit of emission.

The Lens Spectrometer. It is found that many artificially induced radioactive transformations give rise to beta-rays of much greater energy, many spectra extending beyond an equivalent voltage of 10^6 V., and some as high as $5 \cdot 10^6$ V. The semicircular focusing method would require very high or very extensive fields in order to deal with such electrons, and these are difficult to attain. To solve the problem, as well as to achieve higher intensity and resolution of recording at lower velocities, an alternative spectrometer has been developed which utilizes the focusing properties of the short magnetic lens; a solenoidal method has also been employed.

As first devised by Klemperer, the apparatus is as shown in Fig. 133.

S is the beta-ray source, which may be a disk or a slit at one end of an evacuated tube, at the other end of which is situated the aperture of a recording device R . The direct path from S to R is barred by a thick lead stop B , placed at the centre of the magnetic lens M which focuses upon R

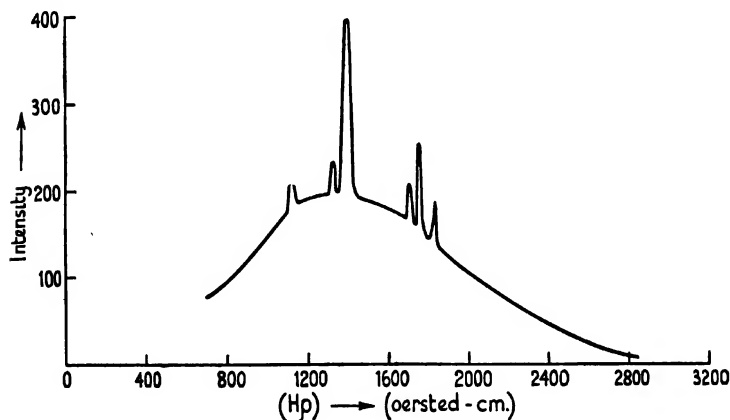


FIG. 132. Beta-ray spectrum of thorium-B.

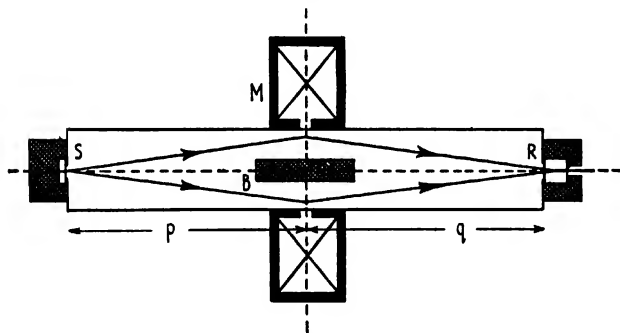


FIG. 133. Lens method of focusing beta-rays.

all electrons incident in it with appropriate velocity. The relation between velocity and focal length has been given earlier (section 31), and, since the lens is short compared with the object-image distance SR , the simple lens formula ($1/f = 1/p + 1/q$) may be used to determine the focal length for a given apparatus. As the current in the coils is increased, the lens will possess this focal length for successively faster groups of electrons and, once the field distribution has been measured, a current-velocity (or momentum $H\rho$) calibration curve may be calculated from IV.12. An independent test of its accuracy may be obtained by employing it to record a well-defined beta-ray spectrum, such as that of thorium-B, which has a very sharp line at $H\rho = 1,390$ gauss-cm. and an upper limit at 3,000

gauss-cm. The recording system may be a Geiger-Müller electrical counter or a Wilson chamber. The instrument may indeed be used either as a direct spectrometer or to select groups of electrons in a required velocity range for observation in the cloud chamber.

In this type of spectrometer only a small part of the electron path needs to be in the magnetic field, and therefore source and recorder can be widely

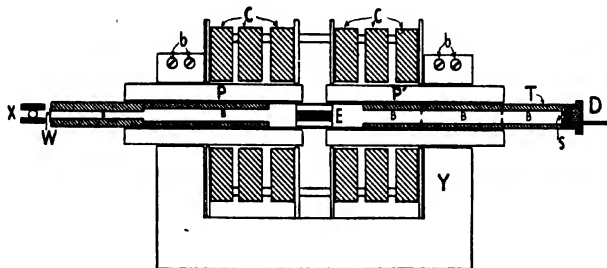


Fig. 134. Magnetic lens with iron yoke, for high-energy beta-rays. (Cosslett.)

separated and more efficiently shielded one from the other than in the semicircular method; also higher energy beta-rays may be focused. A field of 1,000 gauss in an iron-shrouded lens suffices to focus, at a focal length of 17.5 cm., electrons of velocity equal to 0.92 that of light; since $(p+q=4f)$ in the symmetrical arrangement employed, the corresponding length of the focusing tube is 70 cm. Using very heavy iron shielding, Cosslett obtained fields strong enough to focus $15 \cdot 10^6$ V. beta-rays ($v/c = 0.999$) at a focal length of 37.5 cm. The apparatus is in the form of a large yoke-magnet Y (Fig. 134), the poles PP of which are adjustable along the axis and are bored with holes of 5 cm. diameter to accommodate the focusing tube T . A series of lead cylinders BB serve to isolate the recording counter X from the effect of direct and scattered rays from the source S ; the tube is evacuated by a tube through the end-cap. The six strip-wound coils C energize the magnet, and focusing occurs in the gap of a few cm. length between the pole-faces. The maximum field attainable is of the order of 10,000 gauss; the value of $\int H^2 dz$, which determines the focal length (cf. IV.12), is found to be 15 times that given by the unshrouded coils.

The resolving power, $d(H\rho)/H\rho$, of this type of spectrometer may be expressed in terms of the geometry of the system, depending on the size of the ring aperture of the lens, and on the diameters of the source and the recording aperture. In the two models described its value is poorer than that of the semicircular method, but Siegbahn has recently constructed a lens of larger aperture which equals the latter in resolution, and has indicated the direction of further improvement. Deutsch, Elliott, and Evans have made similar experiments, using an unshielded coil, and refined the calculation of resolving power. The chief advantage of the

lens method, in the range of velocities in which the semicircular method may also be used, lies in the higher intensity and dispersion which can be obtained in the recorded beam; weaker parts of beta-spectra may accordingly be investigated with greater accuracy.

65. The Magnetron

The magnetron is a type of thermionic valve used to generate ultra-high-frequency oscillations. Like the original form of beta-ray spectrometer it depends upon the deflecting influence of a transverse magnetic

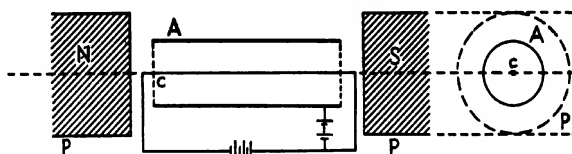


FIG. 135. Magnetron high-frequency valve.

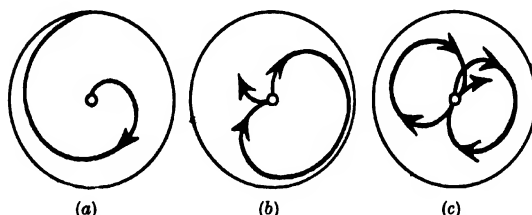


FIG. 136. Electron paths in magnetron. (a) Magnetic field $H < H_c$ the cut-off value; (b) $H \approx H_c$; (c) $H > H_c$.

field, but, as an electrostatic field is also present, it is in principle more closely related to the Slepian electron multiplier (section 53). The electrode system is similar to that of a simple diode (Fig. 135), except that the cylindrical anode A may be split into two or more segments. The filament C is a fine tungsten wire extending along the axis of the anode, relative to which it is at high negative potential ($\sim 1,000$ V.). A strong magnetic field H is applied along the axis of the electrode system by the poles of the electromagnet PP . Electrons emitted normally from the filament, along the lines of the electrostatic field, will also be moving at right angles to the magnetic field and therefore will travel in a cycloidal path towards the anode (cf. section 7). The amplitude of the path will be determined by the field strength and the velocity (\equiv accelerating voltage) of the electron. If the field is sufficiently weak, for a given anode potential, the generating diameter D of the cycloid will be greater than the radius R of the anode and the emitted electrons will reach it, constituting an anode current as in the normal diode. At a critical value H_c of the field, however, the value of D will equal R and electrons will either just reach the anode (Fig. 136a) or graze it and travel a circular path back to the filament (Fig. 136b),

according as their velocity is slightly greater or less than the average given by the energy equality: $eV = \frac{1}{2}mv^2$. When the field is greater than H_c , the anode current will be completely interrupted, the electrons travelling in smaller cycloids back to the filament (Fig. 136c). The corresponding case of crossed electric and magnetic fields between plane parallel electrodes was fully discussed in section 7, leading to a value for the critical magnetic field: $H_c = \sqrt{(2Em/ed)}$, where E is the electric field strength and d the normal separation of the planes. Substituting values for the charge and mass of the electron, and replacing E by V/d , where V is the applied voltage, this becomes

$$H_c = 3.36\sqrt{V/d}. \quad (\text{IX.2})$$

In the magnetron we have an arrangement of concentric cylinders, the filament being a cylindrical wire; a comparable analysis of it has been made by Hull, and extended by Brillouin. The critical cut-off field in this case proves to be

$$H_c = 6.74\sqrt{V/r}, \quad (\text{IX.3})$$

where r is now the radius of the anode.

In practice it is found that the cut-off of current is not sharp, owing

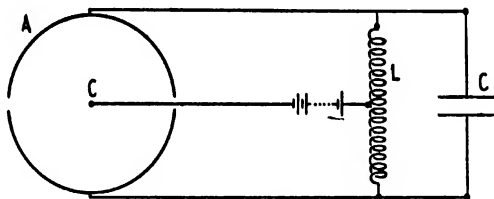


Fig. 137. Split-anode magnetron as oscillator.

to the variation in emission velocity of the electrons, to the voltage drop along the filament, and to departure from true cylindrical symmetry. Consequently, the anode current-voltage characteristic, at constant magnetic field, also does not drop instantaneously to zero at a critical anode voltage, but displays a region of negative slope. Such a 'negative-resistance' or 'dynatron' characteristic indicates that the system can generate self-oscillations, and this may be achieved by connecting a resonant circuit to the two segments of a split-anode magnetron (Fig. 137). Analysis reveals the physical basis of the oscillatory condition: when a potential difference exists between the two segments, the electrons follow paths which carry them ultimately to the segment momentarily at lower potential, so that a negative relation exists between current and voltage. Wave-lengths of less than a metre are readily produced from this system, and values as low as 20 cm. have been reached with water-cooling of the anode. The efficiency is low unless high magnetic fields and small anode diameters are used, when the dissipation of energy from the anode causes difficulty and water-cooling becomes necessary.

Another type of oscillatory condition occurs, however, depending on the transit time of the electron from filament to anode. For the oscillations to be self-maintained the electrons must, on balance, deliver energy to the external resonant circuit. The physical mechanism by which this occurs may be described in outline as follows, supposing the magnetic field to be at cut-off value. If the steady potential of the anode (assumed not to be split) has superimposed on it a small alternating voltage of very high frequency, it will change appreciably in magnitude during the passage of an electron along a grazing path from filament to anode and back. It is convenient to let the frequency be such that the transit time for this path is equal to the period of the applied voltage. Any electron which leaves the cathode at the beginning of a positive half-cycle will experience a greater acceleration towards the anode than an average electron, and may therefore approach close enough to strike it. However, if it only grazes it and returns to the cathode, it will now be moving in a negative half-cycle of the alternating component and will lose less kinetic energy than it gained on its outward trip. It will, therefore, have enough residual energy to carry it to the cathode, to which it gives up this excess energy acquired from the source of alternating voltage.

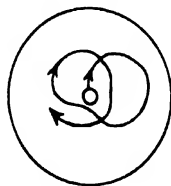


Fig. 138. Precession of electron (decelerated) in magnetron with fluctuating anode voltage.

On the other hand, an electron starting out on a negative half-cycle will suffer less acceleration than the average and will approach less closely to the anode. On the return journey it will likewise lose its kinetic energy more rapidly, the potential of the anode now being higher than before, and will come to rest before reaching the cathode. It thereby delivers energy to the alternating source, and, moreover, can continue to oscillate to and fro (with diminishing amplitude) delivering further minute energy contributions at each excursion until it comes to rest. As it is at the same time precessing about the axis, its path has the form shown in Fig. 138. All electrons leaving the cathode on a positive half-cycle will thus take energy from the alternating source and deliver it to the electrodes, but will be lost at the cathode or anode during their first excursion. Those electrons leaving on a negative half-cycle may make several excursions between the electrodes, in each of which they deliver energy to the external source. Since electrons are leaving the cathode at a constant rate, the net amount of energy delivered over a whole cycle will be a positive quantity.

The action of a split anode magnetron differs in detail from that of the simple type, but the mechanism of energy transfer is essentially the same. If this energy, ultimately derived from the source of steady high potential, is applied to a resonant circuit (cf. Fig. 137) connected across the two anode segments, and if it is sufficient in quantity to make up for the damping losses therein, then self-oscillations will be maintained. The frequency will necessarily be very great, as it must correspond to the transit

time of the electrons between the electrodes; this being of the order of 10^{-9} sec., the frequency must be of the order of 10^8 Mc.p.s. and the wave-length a few decimetres only. Brillouin has shown that the basic wave-length λ is given by

$$\lambda = 15,160/H \text{ cm.},$$

although oscillations are possible in other ranges of frequency above and below this. In such circumstances the values of the inductance and capacity required in the resonant circuit ($f_0 = 1/2\pi\sqrt{LC}$) are so small that a short parallel or coaxial wire system is sufficient. At very high frequencies the resonant circuit may even be provided by a special form of anode assembly itself. Magnetrons of this type have been reported to generate wave-lengths of less than 1 cm.

The transit-time magnetron differs in details of construction as well as in principle from the dynatron type first described. Those electrons which on balance deliver energy to the circuit move a shorter distance on each successive excursion, so that the transit time goes down and the electronic oscillations advance in phase with respect to the basic oscillation of the circuit; eventually they may begin to absorb energy. It is essential to remove them from the system before they come to rest, but, on the other hand, not before they have contributed usefully to the transfer of energy. End-plates are therefore supplied, at an intermediate voltage between cathode and anode, in order to give the electrons an axial drift velocity. Alternatively, the magnetic field may be tilted at a slight angle to the filament, the adjustment being very critical. Okabe found that oscillations could also be maintained if the anode were continuous but the end-plates segmented and connected to the resonant circuit. Good results have also been claimed for a multi-segmented anode and continuous end plates. When one of these methods is adopted to remove the retarded electrons, the transit-time magnetron is found to be much more efficient, as well as of higher frequency, than the dynatron type, and high outputs are possible at very short wave-lengths.

66. *The Cyclotron*

In its essential construction the cyclotron is no more than a two-segment magnetron of large radius and short axial length, with semi-circular end-plates connected to each half of the anode. An external high-frequency voltage is now connected to the two sections (the 'D's') in order to accelerate the electrons (or ions) when they cross the gap between them. In contrast to the magnetron, however, the applied electric field is so low that the particles make a large number of revolutions in the strong magnetic field, following a spiral path from the axis to the anode and gaining energy twice in each revolution. The apparatus is usually employed to obtain high-energy ion beams (protons, deuterons, α -particles) rather than electron beams.

The main features of the apparatus are shown in Fig. 139. The two

anode segments DD , in Lawrence and Livingston's original model almost a metre in diameter and some 15 cm. only in axial extension, contain gas (hydrogen or helium) at a low pressure (10^{-3} – 10^{-4} mm. Hg). A potential difference of 10–100 kV. is applied to the electrodes at a frequency of the order of 10 Mc.p.s. The filament F acts as a source of electrons at the

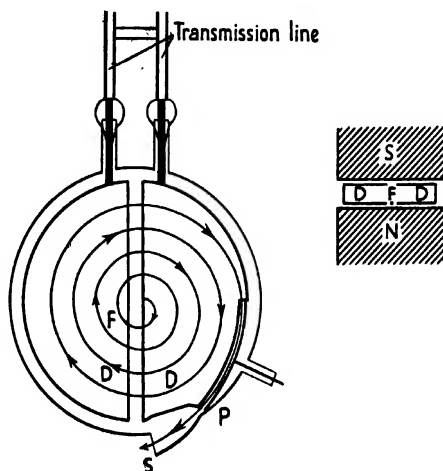


FIG. 139. Cyclotron, in horizontal and vertical section (schematic).

centre of the gap dividing the segments. The emitted electrons produce positive ions in the gas, which are then accelerated by the alternating field towards the segment momentarily at negative potential. Under the influence of the transverse magnetic field, produced by the poles NS , an ion will travel an almost circular path, since the electrostatic field is relatively weak and the shallowness of the segments causes it to be concentrated in the gap (cf. Fig. 140 *a*), the major part of the path thus being in a space where the electric field strength is negligible.

As shown in section 30, the time of revolution in a transverse magnetic field is constant, for a particle of given charge and mass, and independent of the radius of the path, being given by

$$t = 2\pi M/HE, \quad (\text{IX.4})$$

where M is the mass and E the charge of an ion. The frequency of the applied field is adjusted with high accuracy so that its period is equal to t . Hence an ion which crosses the gap at the peak of a negative half-cycle will travel a semicircular path and arrive again at the gap at the instant the negative peak value is reached on the opposite anode segment, and will thus undergo a further acceleration. The process will be repeated at every half-revolution, so that ions rapidly gain energy and consequently travel in a path of ever-increasing radius but of constant orbital time. When they approach the outer wall of the anode, they come under the influence of a

strong extracting field applied to the deflector plate P and pass from the apparatus through the slit S as an ion beam of velocity equivalent to several million electron volts.

Those ions which pass the gap at some other part of a cycle than its peak value will acquire less energy each time and will therefore require to make many more revolutions before they can leave the apparatus. Since the radius of the path is determined only by the ion velocity, in a constant magnetic field, all ions which eventually attain the radius necessary for extraction will have the same velocity, and the emergent beam is homogeneous. It follows from IX.1 that its energy is given by

$$EV = \frac{1}{2}Mv^2 = (EHR)^2/2M,$$

and the equivalent voltage

$$V = (E/2M)(HR)^2. \quad (\text{IX.5})$$

Thus the final voltage is determined by the strength of the magnetic field and by the maximum radius R of the anode segments. It is limited by the practical consideration of the large magnet poles and correspondingly massive iron yoke required. The value of R is limited also by the fact that the capacity of the anode (dependent on its size) forms the largest part of the impedance in the resonant circuit of the high-frequency oscillator, and thus determines the shortest wave-length which may be generated. The minimum period of revolution of the ions is thus fixed, and equation IX.4 at once gives the value of H required; when R is of the order of decimetres, H must be of the order of 10^4 gauss, giving an emergent beam of some millions of electron volts in energy. The necessary wave-length is 10–20 metres, or a frequency of 30–15 Mc.p.s. In later models generating very high energies a quarter-wave transmission line is used to feed the energy to the resonant circuit which is formed by the segments themselves.

The number of revolutions made by an ion travelling across the gap at the peak of the applied voltage, V_p , is equal to the final voltage V divided by $2V_p$, since it receives two impulses per revolution. In a proton beam of 2 MeV. energy such an ion will have made 100 revolutions under an applied potential of 10 kV., and those emitted at other parts of the cycle correspondingly more. For ions making more than about 150 revolutions the chance of collision with a gas molecule is so high that they are mostly lost from the beam. Therefore only ions emitted near the peak of the voltage wave can appear in the final beam, which thus has a pulse character. The pulse frequency is so high, however, that the beam is effectively continuous. The intensity of the emitted beam is determined otherwise by the emission from the filament and by the gas pressure, and is of the order of 50 microamperes. With the use of an arc source of ions Livingston has obtained beam currents as high as 300 microamperes.

The intensity of the initial beam at the centre of the apparatus is conserved by two focusing properties of the fields: an electrostatic effect at

small values of the radius, and a magnetic effect at large radius. In the gap between the anode segments the electric field has the configuration shown in Fig. 140*a*, the lines of force (broken lines) running from edge to edge and the equipotentials intersecting them orthogonally. The gap constitutes a true cylindrical lens for the ions, entirely similar to the two-tube lens for electrons (section 24); the positive ions, travelling in reverse

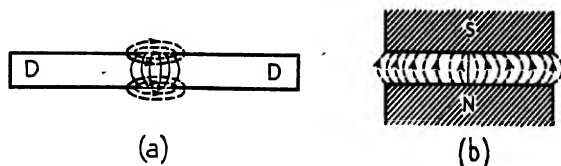


FIG. 140. Cyclotron: focusing effect of (a) electrostatic field in gap between 'D's'; (b) radial component of magnetic field.

direction to electrons, will experience a strong focusing action when passing across the gap into the negative segment. Twice in every revolution, therefore, errant ions will be deviated towards the 'axis', which is here the median plane of the anode, and will thus be concentrated into a flat beam. The focal length of the lens, however, is proportional to the voltage of the beam (cf. III.23), and so the focusing effect is appreciable only during the early part of the path and decreases with radius. The influence of the fringing field of the magnet then comes into play. The curvature of the magnetic lines (Fig. 140*b*) introduces a horizontal component into the field, which is directed inwardly above, and outwardly below, the median plane. Consequently ions in a plane above or below the latter will experience a force directing them towards the median plane (Left-hand Rule). As the magnetic field is only truly normal to the poles near the centre of their faces, this effect is operative over most of the path, and the beam is found to be well concentrated on emergence, having a cross-sectional area of about 1 cm.²

The above discussion has assumed the mass of the accelerated particles to be constant, but at the high velocities finally reached the relativistic correction becomes appreciable: a 100 MeV. deuteron would have a mass 10 per cent. greater than its rest mass. Hence, from IX.4, the strength of the magnetic field should increase correspondingly with the radius, since the time of revolution is fixed by the oscillating source. When this is achieved by proper shaping of the poles, it is found that the focusing effect of the fringing field is lost, and a defocusing action occurs instead. In present cyclotrons this sets a limit of about 16 MeV. to the attainable energy of a deuteron beam, which can only be overcome by using very high voltages between the segments, in order to increase the electrostatic focusing action. Such a voltage is in any case desirable in order to reduce the number of revolutions made by an ion, and thus to increase the beam intensity. The new Berkeley cyclotron employs a peak voltage in the

neighbourhood of 10^6 V., and is expected to yield beams of 100 MeV. energy.

67. *The Betatron, or Induction Accelerator*

The relativistic increase in mass which, as just noted, affects the production of high-energy ion beams in the cyclotron, is much more important for electrons with their bigger ratio of charge to mass. It is, indeed, so large that the velocity of electrons becomes almost constant when the accelerating potential exceeds a few million volts: at $3 \cdot 10^6$ V. it is 0.99, and at 10^7 V. it is 0.999 that of light. Therefore it is in principle impossible to establish resonance between the circulating electron and the applied electric field, even if the very high radio frequencies could be generated which are demanded by the relatively large charge-mass ratio (cf. IX.4). The cyclotron thus cannot be employed to generate high-voltage electron beams: and only very energetic beams (50–100 MeV.) are of value for work on atomic disintegration. No alternative method existed until the development of the betatron.

It is a fundamental law of electromagnetic induction that a change of flux linking a circuit gives rise to an electromotive force in that circuit of magnitude dependent on the rate of change of flux. The induction is none the less effective when the circuit is an electron moving in a circular orbit of radius r . If ψ be the magnetic flux enclosed by its path, then the induced electric field E (in electromagnetic units) will be

$$E = \dot{\psi}/2\pi r;$$

and the rate of change of momentum, by Newton's Second Law, is

$$\dot{p} = eE = e\dot{\psi}/2\pi r, \quad (\text{IX.6})$$

writing p for the momentum of the electron. We assume for the moment that the electron can be constrained in this orbit of constant radius during the change of flux. Then, if the flux changes from ψ_0 to ψ_1 in a time interval t , the resulting change in momentum will be

$$p_1 - p_0 = e(\psi_1 - \psi_0)/2\pi r.$$

The final momentum of an electron starting from rest will then be

$$p_n = \frac{e\psi_n}{2\pi r} \quad (\text{IX.7})$$

if ψ_n is the maximum value of the flux, the initial value being zero. But, by equation IX.1, the momentum is also given by

$$p = mv = erH_r,$$

where H_r is the field strength at the orbit. Hence $\frac{\psi_n}{\pi r^2} = 2H_r$, or the average field strength across the orbit must be equal to twice the value at the orbit. If this condition is satisfied at every instant of the variation of flux, the electron will be maintained in a path of constant radius r . The correct shape of the poles to give a field of the required distribution is a

technical problem easily solved: they must have the shallow conical form shown below in Fig. 142 (P, P).

If high final energy is to be imparted to the electron, the mean field within the orbit must rise to a high value in a very short time. If it rises linearly from zero to 10,000 gauss in $1/1,000$ sec., and the path radius is 10 cm., then from IX.6 the electric field strength will be

$$E = \dot{H}r/2 = 0.5 \text{ volt/cm.}$$

From IX.7 we have for the resulting velocity

$$v = \frac{eHr}{2m} = (e/2m)10^5 \text{ cm. sec.}^{-1},$$

if the electron starts from rest. Insertion of the value for e/m_0 ($1.76 \cdot 10^7$ e.m.u./gm.) gives a velocity of nearly 10^{12} cm. per sec., considerably beyond the velocity of light. If correction is made for the increase in mass of the electron with velocity, it still appears that the electron is accelerated to approximately the velocity of light in a very short time compared with $1/1,000$ sec. Hence it will travel in this time a distance of the order of 3×10^7 cm. in an electric field of 0.5 V. per cm. Its final energy must therefore be about 15 MeV., being primarily determined by the strength of the induced electric field in which it moves. Provided that the initial acceleration is sufficiently great, the period of the flux change is not important, since it affects in opposite senses the electric field strength and the distance travelled in it. The beam energy W therefore depends on the total flux change (IX.6) and thus on the final value of the field and the radius of the orbit. It may be shown that

$$W = 3 \times 10^{-4} H_r r - 0.51 \text{ (MeV.)},$$

where H_r is the maximum field strength at the orbit. In the example given here, H_r is 5,000 gauss, being half the average over the whole orbit, and $r = 10$ cm. so that $W = 14.5$ MeV., in agreement with the approximate value deduced above. The original betatron used an orbital field of 3,000 gauss and a radius of 5 cm., generating a beam of 2.3 MeV. electrons.

The construction of such an electron accelerator was attempted by Walton in Cambridge and by Wideröe in Germany, but the difficulties were only finally solved by Kerst (of Illinois) after a careful analysis of the conditions necessary for stability of the electron in its orbit. The radius of the orbit will remain constant during the change of flux, so long as the field at the orbit is always half the average value over the area enclosed by the orbit, as shown above; but the form of the field should be such that electrons scattered into larger or smaller orbits are returned to the equilibrium orbit. It is also essential to the production of a useful beam that the electrons should be constrained to rotate in a given plane: any deviation of the path from this plane orbit should bring into play restoring forces of adequate magnitude. Kerst determined the nature of these 'focusing' conditions, which must be rigidly satisfied, since an electron will make some 10^5 – 10^6 revolutions in the field during the rise in flux. As regards

deviation from the median plane, the stabilizing condition is the same as in the cyclotron: the magnetic field must have a horizontal component of opposite sign on either side of this plane. Such a bulging of the lines of force occurs naturally with the type of conical pole-piece needed to reduce the field to the correct value at the orbit. Deviated electrons will then execute a damped oscillation about the orbit, returning to the median plane after a few revolutions.

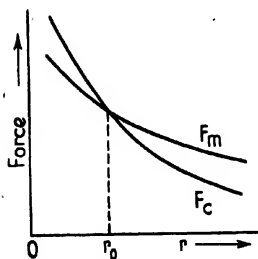


FIG. 141. Betatron: radial variation of magnetic (F_m) and centripetal (F_c) force on circulating electron.

Radial deviation from the orbit will give rise to a similar oscillation in the median plane itself, provided that the radial decrease in field strength is less than a given value. The centripetal force F_c acting on an electron moving at velocity v in an orbit of radius r is mv^2/r , decreasing in magnitude inversely as the radius. It is supplied by the action of the magnetic field, exerting a force $F_m = eHv$. In the equilibrium orbit these two forces are necessarily equal: $F_c = F_m$ (cf. Fig. 141). If the field decreases less rapidly than $1/r$, then at larger orbits more than the required centripetal force will be applied to the electron, which will be directed back to the equilibrium orbit; if it now passes momentarily to an orbit for which $r < r_0$, then the field exerts less than the necessary force and the electron moves outwards again. It oscillates about the equilibrium orbit, the amplitude decreasing because of the rising value of H ; it is, in fact, inversely proportional to the root of H . The radial oscillations are thus also rapidly damped, and the orbit of radius r_0 is maintained so long as the field is increasing. In practice the pole-pieces are so shaped that the magnetic field decreases proportionally to r^{-1} .

As constructed by Kerst the betatron has the form indicated in Fig. 142. The conical pole-pieces P, P are surrounded by an evacuated glass tube of annular form, the cross-section being that of a truncated cone; with reference to American conditions, it is described as of 'dough-nut shape'. The electron injector consists of a filament F , concentrating grid G , and anode A , to which an injection voltage of about 100 V. is applied. The magnet is comparatively small and is energized by a 600 c.p.s. generator, the coils forming the inductance of a resonant circuit; a large bank of condensers is required across the coils in view of the low frequency of operation. The first quarter-cycle of the sinusoidal wave is used for accelerating the electrons, the injector supplying a beam at the instant $\theta = 0$ and the final beam being extracted at $\theta = \frac{1}{2}\pi$; that is, the flux change occupies $1/2,400$ sec. The maximum value of the field at the orbit is 3,000 gauss, giving a voltage gain per revolution of about 12 V. at a radius of 5 cm.

The beam may be extracted in a number of ways, that depicted in Fig. 142a utilizing iron dust in the pole caps P , which therefore become

saturated more readily than the main body of the poles, so that towards the end of each quarter-cycle the field at the orbit increases faster than is necessary and the beam spirals in and strikes the target at T . Alternatively, solid pole caps may be used and extra windings employed to weaken the field at the orbit, and thus allow the beam to increase in radius and strike the rear of the injector (Fig. 142*b*). In either case X-rays of

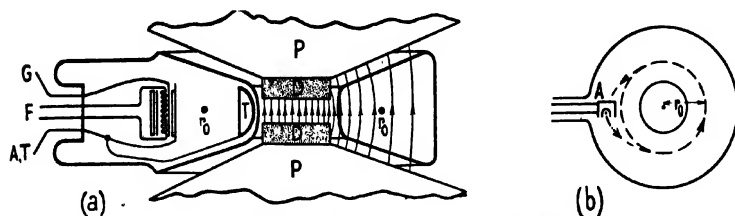


FIG. 142. Betatron: (a) vertical, (b) horizontal, section.

extremely short wave-length are generated; in this original apparatus the radiation corresponded to an electronic energy of 2.3 kV., the beam current being about 0.5 microamp. A later model operates at 180 c.p.s., having a magnet of pole diameter 48 cm. and yoke 90 cm. \times 150 cm., weighing 3.5 tons (as compared with the 60 tons of the comparable cyclotron). The equilibrium orbit is of 7.5 cm. radius and the final beam has an energy of 20 MeV. and a current of 1 microamp. At this potential the efficiency of X-ray production is very high, about two-thirds of the energy appearing as radiation. The heating effect is therefore not great, and, the focus being excellent, the beam spot is very small and the dissipation of heat in the target rapid, so that the latter is not melted. In this model it is possible to make use of the scattered electrons from the target, as well as the X-rays, for radioactive and therapeutic investigations.

Whilst the magnetic field is changing, the beam is continuously accelerated and thus no phase relation between electron and field is involved. The relativistic change in mass consequently does not limit the use of the betatron as it does the cyclotron. The magnetic fields required are neither extensive nor very great in magnitude, and there is no major obstacle to the construction of models to operate at yet higher voltages. The problem becomes mainly the engineering one of supplying adequate electric power at the required frequency. A 100 MeV. apparatus is under construction in America, and Cockcroft has discussed a project for the production of a 240 MeV. beam, at which energy it should be possible to reproduce some of the cosmic-ray phenomena.

68. Radio Valves

In most of the electron optical problems discussed in earlier chapters the field distribution has displayed rotational symmetry, being the same in all planes containing the axis. The applications so far described in this chapter, on the other hand, mostly employ magnetic fields of cylindrical

symmetry: the distribution is the same in all parallel planes transverse to the axis, the field at a given radial distance from it being independent of the axial coordinate z . The normal types of thermionic valve similarly utilize electrostatic fields of cylindrical symmetry, although the symmetry is rarely complete. A certain degree of focusing is displayed, being small in an ordinary triode and pronounced in a beam tetrode, the field in each

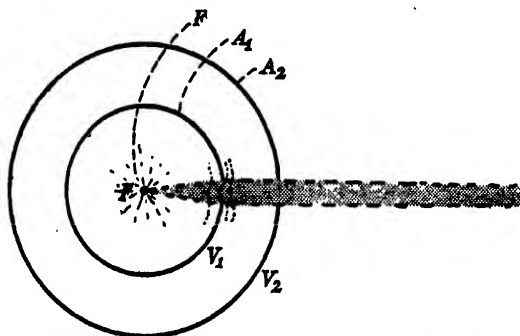


Fig. 143. Cylindrical lens formed between longitudinal slits in two concentric cylinders ($V_2 > V_1$).

case acting as a cylindrical lens in the glass optical sense (as distinguished from the two- or three-cylinder electrostatic lenses, which display rotational symmetry).

The simplest form of cylindrical lens is formed by two parallel longitudinal slits in two concentric cylinders A_1 , A_2 maintained at different positive potentials V_1 , V_2 , with respect to an axial filament F at potential V_0 (Fig. 143). Provided that the cylinder radii are large compared with the width of the slits, the field has a similar form to that between two plane electrodes with circular apertures (cf. Fig. 37), the cylinders being effectively parallel planes. The system behaves in every way like an aperture lens, and Davisson and Calbick have shown that the focal length is given by

$$f = 2V_1/(E_1 - E_2), \quad (\text{IX.8})$$

where E_1 and E_2 are the field strengths on either side of the slit in the smaller cylinder. The circular aperture lens has a focal length given by an identical expression (III.17), save that the factor 4 appears in place of 2. Hence the cylindrical lens is twice as powerful as its circular counterpart, as might be expected from the deeper penetration of the stronger field into the weaker, through the longitudinal slit. Davisson and Calbick determined the focal length by finding the potentials necessary to give a parallel emergent beam, in which case it must be equal to the radius of the inner cylinder.

At the other extreme lies the case of a slit which is so wide that the inner cylinder shrinks to two narrow plates or rods (Fig. 144 a). The curvature of the field lines is then little altered from the simple cylindrical form, so

long as the inner electrode G (or 'grid') is at positive potential with respect to the filament F , and therefore the divergent beam emitted from the latter experiences a very small focusing effect. A negative potential on the grid, however, radically alters the field distribution, negative equipotentials surrounding its two rods and giving rise to a more pronounced

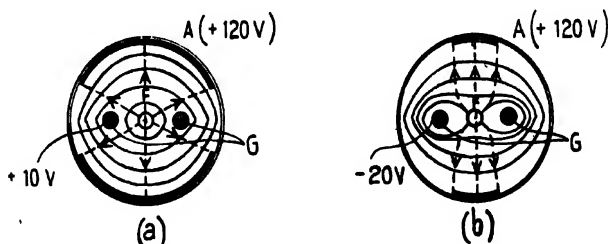


FIG. 144. Triode valve: focusing action of grid. (a) $V_G = +10$ V.; (b) $V_G = -20$ V.

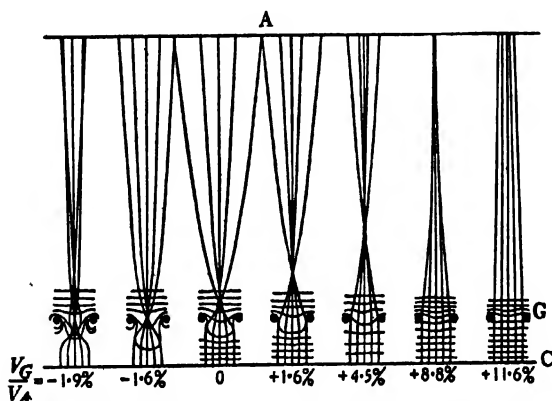


FIG. 145. Focusing action of grid in triode valve with plane electrodes, for various ratios of V_G/V_A (after Knoll).

focusing effect the more negative they are made (Fig. 144 b). The similar effect in a plane parallel arrangement of cathode, grid, and anode is shown in Fig. 145 for various grid potentials. When the grid is at such a positive voltage that the field strength has almost the same value on either side; the beam suffers negligible concentration; but decreasing the voltage, and finally making it negative, gives rise to an increasing focusing effect. When the cross-over occurs close to the grid, the beam leaves it with greater divergence than from a positive grid. At a high negative potential, however, a second cross-over may occur and the beam meets the anode in a comparatively small spot.

The construction of a triode valve is, of course, more complicated than either the plane parallel system or that of Fig. 144. The grid is not

composed of two parallel rods but of a flattened helix of small pitch. It may be approximately treated as a rod system on which is superimposed a set of parallel conducting loops at regular axial spacing, so that true cylindrical symmetry is lost. In some degree the latter act as weak aperture lenses and concentrate the emission into the transverse planes intermediate between the grid loops. The normal valve shows very little of

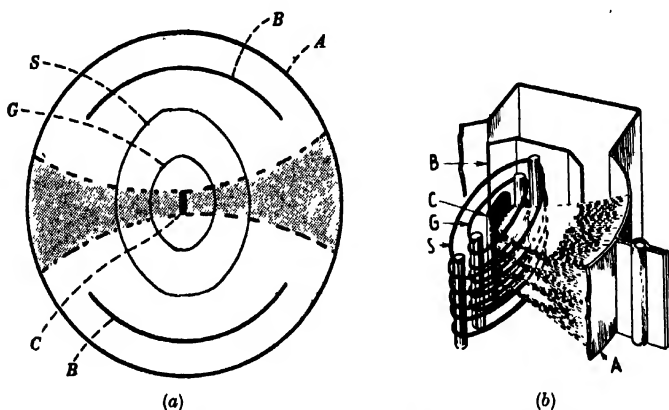


FIG. 146. Beam power valve: (a) horizontal section; (b) electrode form.

this effect, no concentration being desired and therefore the necessary alinement, spacing, and potential ratio of the grids not being provided. Some types of tuning indicator (or 'magic-eye'), however, utilize this variation of beam size with grid voltage, the latter being controlled by the magnitude of the signal and thus by the accuracy of tuning.

The beam type of power valve, on the other hand, is directly based upon the principles of cylindrical focusing. It is designed to fulfil the same function as a pentode, and therefore must effectively suppress the secondary electrons emitted from the anode under the influence of the screen grid. The construction is shown in Fig. 146, where *C* is the rod cathode, *G* the control grid, *S* the screen grid, and *A* the anode. Two segmental electrodes, *B, B*, the beam-forming plates, are interposed between the screen and the anode. The system is usually described as a 'beam tetrode', but it is in actuality a pentode in which one electrode is much less extensive than the others. The beam-forming plates are at cathode potential and act in essentially the same way as the rod grid of Fig. 144, forming the cathode emission into fan-shaped beams which impinge on the anode over a sharply defined area only. At the same time, the accurately alined grids concentrate the emission between consecutive loops, so that the beam consists of nearly parallel sheets and not of a continuous distribution. These wedges of electrons are retarded in the space between the beam-forming plates, since they are at cathode potential, and set up a space charge before the anode which effectively prevents secondary elec-

trons from reaching the screen; this action is assisted by the large screen-anode distance. The accurate alinement of the two grids also ensures that very few electrons strike the screen, the directed beams having very small axial thickness when passing between the loops. The screen current is therefore reduced to less than 2 per cent. of the anode current. The anode characteristic of the beam valve is found to have a sharper top bend than a pentode, whilst the straight part extends over a wider range of anode voltage; as a result its power output and efficiency are greater for a given amount of harmonic generation. The mutual conductance also proves to be several times greater than that of a pentode, and has led to its use in wide-band amplifiers for television reception.

69. *Mass Spectrographs*

Electrically charged particles may be separated according to their charge-mass ratio by deviation in electromagnetic fields, on similar lines to the separation of electrons according to velocity of ejection in beta-ray spectrometry (section 64). The particles to be separated and identified are usually isotopes of a given element: they have differing nuclear masses (atomic weights) but identical electron configurations (atomic number), and hence cannot be chemically separated. Any beta-ray spectrometer may, in principle, be used as a mass spectrograph for isotope separation, a source of positive ions ('positive rays') taking the place of the β -radio-active body; it is an advantage to accelerate all ions to a common velocity before they enter the dispersing fields. Analysis of particles of given velocity but differing values of E/M then takes the place of velocity analysis of beta-rays of uniform charge and mass.

From the ray-path equations (II.16 and IV.8) it is clear that particles accelerated by a common voltage have identical trajectories in pure electrostatic fields whatever their E/M ratio, but have different paths in pure magnetic fields. Hence a combination of electric and magnetic fields is required in any practical mass spectrograph, an electrostatic field being indispensable for producing and collimating the ion beam. A considerable variety of systems has been employed for the purpose, and only those of major historical and practical interest will be described here.

(a) *J. J. Thomson's Method.* The original method used by Thomson for positive ray analysis consisted essentially of the system of concurrent electric and magnetic fields discussed in sections 5 and 30. When both fields act in the same direction, the trajectory of a charged particle is a helix of increasing pitch about the field axis: the radius depends on the strength of the magnetic field H , and the pitch on that of the electrostatic field E (cf. Fig. 4). In J. J. Thomson's method the positive ions were produced in a cold cathode discharge and passed, with a wide distribution of velocities, through a hole in the cathode to enter an electromagnetic field normal to its axis. The fields being small in extent and in magnitude, only a slight deviation of the particles results before they meet a screen or photographic plate (Fig. 147); i.e., only a small portion of one turn of

the helical path is traversed. Curved traces are made on the screen by particles of common E/M and different velocity.

From the discussion of the deflexion of electron beams by charged plates or magnetic coils (section 56) it follows that the electrostatic deflexion suffered is inversely proportional to the square of the velocity

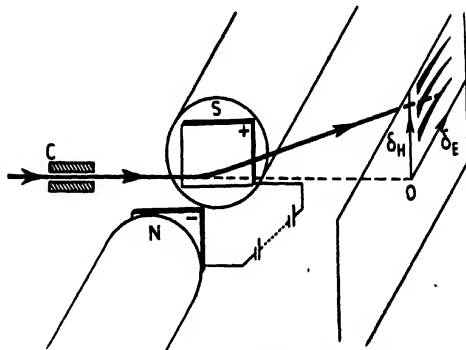


FIG. 147. Concurrent electric and magnetic fields as mass spectrograph (schematic).

($\delta_E \propto V^{-1} \propto v^{-2}$), whereas the magnetic deflexion is inversely proportional to the velocity (cf. equations VII.5 and VII.8). Hence, particles of given E/M suffer a net deflexion which gives a parabolic trace on a plane normal to their initial direction; the vertex of the parabola is at the origin, corresponding to infinite velocity, the slower particles falling farther away along the trace. From a mixed source of ions a set of parabolas will be formed, corresponding to the different ratios of E/M present in the beam. As the ratio of the electric to the magnetic deflexion is proportional to the velocity, particles of the same velocity but differing E/M will fall on straight lines radiating from the origin. By comparison with the traces given by pure ion sources in the same fields, the ratio of charge to mass for the different components in a mixed beam may be found and isotopes of different mass identified. Thomson's original method had a resolution corresponding to a mass difference of about 5 per cent.

(b) *Aston's Mass Spectrograph.* The resolution was greatly increased by Aston's method, in which electric and magnetic fields crossed at right angles (cf. section 6) are applied in succession to deflect the ion beam, instead of the parallel coincident fields of Thomson's method. It follows from equations VII.5 and VII.8 that the dispersion produced by an electric field is twice that of a magnetic field: $d\theta/\theta = 2d\phi/\phi$. Hence, if the deviation ϕ produced by the latter is equal to twice that of the former θ , and in the opposite sense, the emergent beam will show parallel paths for particles of identical E/M but differing velocities (direction HG , Fig. 148). If, however, $\phi > 2\theta$, the beam will converge towards a point, and if the magnetic field is relatively so strong that $\phi = 4\theta$, the ion beam will leave it with a degree of convergence equal to the divergence introduced by the

electric field earlier in the path. All particles of the same E/M will thus be brought to a focus at a point F given by $EH = HF$, EF being parallel to HG . The net deviation suffered by the beam is θ , which depends on the ratio E/M . So long as θ remains small, particles having a range of values of E/M about that corresponding to the focus at F will give a series

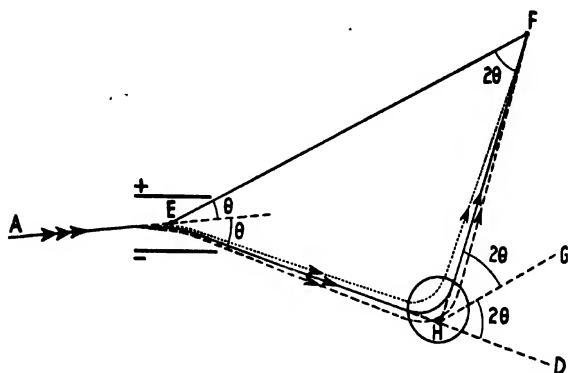


Fig. 148. Crossed electric and magnetic fields as mass spectrograph (schematic.)

of sufficiently sharp images on a plate placed along the direction EF . Isotopes having masses in the neighbourhood of that of a reference element may then be identified.

Since Aston's method brings to the same focus all ions of common E/M whatever their velocity, it has both higher resolving power as to mass and greater sensitivity as to abundance than Thomson's method. He was able to distinguish ions differing in mass by only one part in 600, and to compare the ratio of two masses to one part in 10,000.

(c) *Semicircular Method.* The semicircular method of analysing beta-rays in a uniform transverse magnetic field, described in section 64, may be applied also to mass spectrography. From equation IX.1 it follows that ions of differing mass will be brought to different focuses in a given field (cf. Fig. 131). In order to obtain sharp images it is necessary to employ ion sources giving an effectively monochromatic beam. This was first done by Dempster, using the electrons from a thermionic filament to ionize the gas or vapour containing the elements under investigation; Thomson and Aston had used cold cathode discharges, which give ion beams with a wide range of velocities. Dempster's apparatus is essentially the same as that of Fig. 131, with the ion source and accelerating electrodes replacing the beta-ray source. The recording system may be an ionization chamber or Geiger counter instead of a photographic plate. The former are particularly suited to measurements of relative abundance; ions of different E/M may be focused into their aperture by a change in the strength either of the magnetic focusing field or of the electric field used to accelerate the ions. The sensitivity of the method has been increased

by both Dempster and Bainbridge by the introduction of a collimator, or velocity selector, in the form of a radial electrostatic field between curved condenser plates. It is then possible to distinguish between ions which differ in mass by only one part in 4,000.

(d) *Other Methods.* A number of other systems have been employed with varying success for mass spectrography or for isotope separation on a larger scale. The principles of both the magnetron (cf. equation IX.3) and the klystron (see next chapter) have been applied to the separation of the uranium isotopes required for the production of atomic energy, but with less success than the semicircular method ('calutron'). The use of the short or long magnetic lens (sections 30 and 31) is theoretically possible, as in beta-ray spectrometry, but it is difficult at present to obtain both high resolution and quantity sensitivity owing to the effect of aberrations. Other systems, in the form of specially shaped fields, have been proposed from time to time for bringing a parallel ion beam to a point focus at positions varying with E/M . A valuable summary of the various methods is given in the paper of Jordan and Young quoted below.

FURTHER READING

1. Beta-ray Spectrometer:

Rutherford, Chadwick, and Ellis, *Radiations from Radioactive Substances* (1930).
Klemperer, *Phil. Mag.* **20**, 545, 1935.
Coslett, *Proc. Phys. Soc.* **52**, 511, 1940; *J. Sci. Inst.* **17**, 259, 1940.
Siegbahn, *Ark. Mat. Astr. Fys.* **28**, 17, 1942; and **30**, 1, 1943 (in English).
Deutsch, Elliott, and Evans, *Rev. Sci. Inst.* **15**, 178, 1944.

2. Magnetron:

Hull, *Phys. Rev.* **18**, 31, 1921.
Megaw, *G.E.C. Journal*, **7**, 94, 1936; *J.I.E.E.* **72**, 326, 1933.
Gill and Britton, *J.I.E.E.* **78**, 461, 1936.
Pidduck, *Q. J. Math.* **7**, 201, 210, and 241, 1936; **8**, 75, 1937; **9**, 221, 1938.
Brillouin, *Phys. Rev.* **60**, 385, 1941.
Harvey, *High Frequency Thermionic Tubes* (1944).
Brainerd, *Ultra-high Frequency Techniques* (1943).
Gabor, *Proc. Roy. Soc. A*, **183**, 436, 1945.

3. Cyclotron:

Lawrence and Livingston, *Phys. Rev.* **38**, 834, 1931; **40**, 19, 1932.
Cockcroft, *J. Sci. Inst.* **21**, 189, 1944.

4. Betatron:

Kerst, *Phys. Rev.* **60**, 47, 1941; **61**, 93, 1942; *Rev. Sci. Inst.* **13**, 387, 1942.
Cockcroft, loc. cit.

5. Radio Valves:

Knoll, *Z. techn. Phys.* **15**, 584, 1934.
Thompson (H. C.), *P.I.R.E.* **24**, 1276, 1936.

6. Beam Valves:

Shoenberg, Bull, and Rodda, Brit. Patent No. 423932/1933-5.
Rothe and Kleen, *Beiträge zur Elektronenoptik*, p. 94 (1937).
Schade, *P.I.R.E.* **26**, 137, 1938.

7. Mass Spectrographs:

J. J. Thomson, *Phil. Mag.* **13**, 561, 1907; **21**, 225, 1911.

Aston, *Phil. Mag.* **38**, 709, 1919; and *Mass-Spectra and Isotopes*, Longmans, Green & Co. (1933).

Dempster, *Phys. Rev.* **11**, 316, 1918; **20**, 631, 1922; **33**, 1019, 1929.

Bainbridge, *J. Franklin Inst.* **215**, 509, 1933.

Bainbridge and Jordan, *Phys. Rev.* **50**, 282, 1936.

Jordan and Young, *J. Appl. Phys.* **13**, 526, 1942.

Smyth, *Atomic Energy*, chaps. ix and xi (U.S. Government Printing Office, and H.M. Stationery Office), 1945.

CHAPTER. X

VELOCITY-MODULATED BEAMS

70. Principles

THE electron differs fundamentally from a light wave in that, being a charged particle, it may be subjected to acceleration or retardation by the application of an electric field. When an electron beam enters a steady field, the same acceleration (apart from relativity effects) is experienced by each of the component electrons; the beam is affected as a whole, and continues with a uniform (but new) charge distribution along its length after leaving the field. When, however, a rapidly fluctuating field is applied to a beam which is initially homogeneous as to velocity, the result is very different. If the field fluctuates in a sinusoidal manner over a sharply defined region only, on either side of which the beam moves in field-free space, then electrons entering during a negative half-cycle will suffer retardation, whilst succeeding electrons experience an acceleration during the positive half-cycle. The 'packet' of electrons which passes through the field during a complete cycle will then contract along the axis of the beam as it continues in field-free space, the foremost members, of reduced velocity, being overtaken by those which experienced a positive acceleration. After a certain distance a sharp pulse of current will have been established, so that the originally continuous beam becomes a succession of groups or 'bunches' of electrons, which change position in a regular manner with distance as the beam proceeds, somewhat after the nature of rarefactions and compressions in a sound wave. The applied field, or 'buncher', in effect imposes longitudinal oscillations of density on the electron beam, which consequently differs in character from the flow of electrons in a normal thermionic valve, where variations in grid voltage produce current pulses of coherent form.

Oscillations thus set up do not suffer from the limitations of power loss and frequency that beset those generated in material resonant circuits containing inductance and capacity. They can be used to convey oscillatory energy to a quarter-wave transmission line or a cavity resonator, and with very high efficiency, since the input power absorbed at the buncher is almost negligible. The system can also be used as an amplifier, when the energy exchange, or 'reaction', falls short of the value needed to maintain oscillations in the external circuit. Frequencies higher than 10^8 Mc.p.s., or wave-lengths of a few centimetres, may be readily generated by this means with an efficiency greater than is attainable with the magnetron. As the determining factor is the relative axial position or phase-relation of the electrons, the method is often referred to as 'phase-focusing', or alternatively as 'space-time focusing', and the practical devices as 'velocity-modulation' tubes. The term 'focusing' is here used in the optical sense of concentrating at a point an appreciable proportion

of the emission but this is effected intermittently only and no focusing occurs in the wider sense of gathering the radiation incident on an aperture and producing an image, whether of a point or extended object. The term 'concentration' or 'density-modulation', rather than focusing, would fit the case better. No similar phenomenon is found in light optics, as it could only be produced by periodical variation of the refractive index of the transmitting medium. The term 'phase-focusing' is used in quite another sense in light microscopy, providing an additional reason for avoiding its use to describe a method of modulating an electron beam.

Pulse Formation. In order to appreciate in detail the bunching process, suppose a beam of electrons to be generated in a three-element gun (Fig. 149), under an accelerating potential V_0 . It passes through two closely

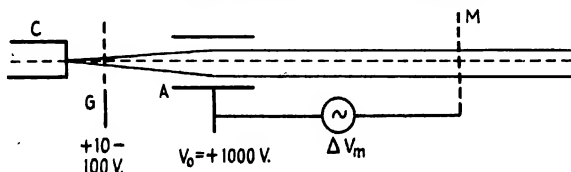


FIG. 149. Electrode system for velocity modulation of electron beam (schematic).

spaced modulating grids M (the 'buncher') to which an alternating voltage ΔV_m is applied at such a high frequency that the time of transit of the electron between the grids is comparable with its period. The field strength must vary appreciably during the passage of a single electron, but not so rapidly as to reverse before it passes the further grid; in practice the transit time is adjusted to be less than half the period T of the oscillator. Successive electrons approaching the first grid may be assumed to have the same velocity v_0 and to be equally spaced along the beam, as is indicated in Fig. 150 by a regular row of points passing the buncher. Their subsequent motion is shown by a set of distance-time curves, the slope of each being given by the emergent velocity of the corresponding electron as determined by the instantaneous strength of the electric field, varying sinusoidally as depicted; such a representation of the relative separation of electrons along the length of the beam is known as an Applegate diagram.

An electron A entering the field at its zero value ($t = 0$) is uninfluenced and continues with the same velocity v_0 . Those immediately succeeding it, up to B , suffer increasing retardation and leave with less velocity, B having a minimum value $v_0 - \frac{1}{2}\Delta v_m$, where Δv_m is given by the energy relation

$$e\Delta V_m = e(V_0 + \Delta V_m) - eV_0 = \frac{1}{2}m(v^2 - v_0^2) = \frac{1}{2}m\Delta v_m(v_1 + v_0) \\ \approx mv_0\Delta v_m, \text{ if } \Delta V_m \ll V_0.$$

Up to time $\frac{1}{2}T$, therefore, the space-time lines have decreasing slope, indicating that these electrons increasingly lag behind each other and that the beam thins out with time. Between B and C the retardation decreases and the corresponding lines increase in slope up to the point C , which

suffers no change and continues with velocity v_0 . Subsequent electrons are accelerated and the space-time lines continue to steepen, up to a maximum at D , with $v_d = v_0 + \Delta v_m$; after D the slope decreases again through E to a second minimum at F , the peak of the next negative half-cycle. As a consequence, the electrons on either side of C approach it as

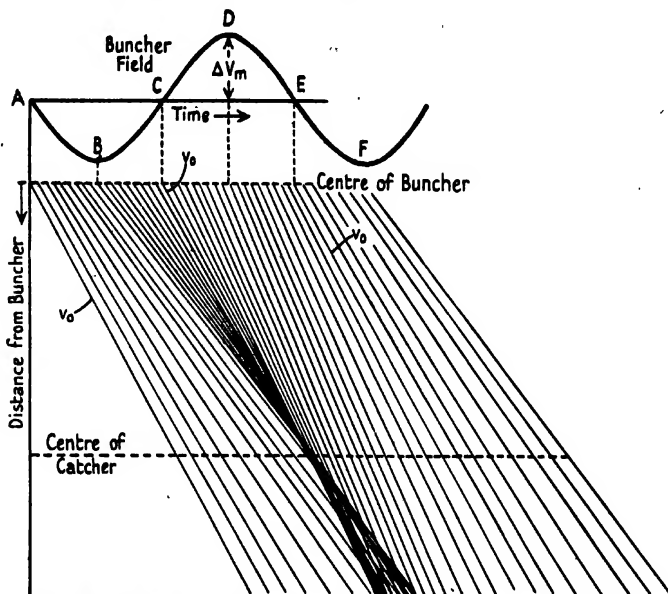


Fig. 150. Distance-time chart for electrons passing in succession through a modulating electrode or buncher (Applegate diagram).

time goes on, those in front of it being slower, and those behind faster, than v_0 ; so that after a certain time t_x the electron density about C reaches a maximum, as the Applegate diagram shows by the corresponding density of lines. Beyond this the faster electrons overtake C and the slower lag behind it, and after a further interval of $T/2$ the electron density about C will become a minimum. Similarly those electrons (A and E) which pass the buncher at the middle point of a falling wave of voltage experience at first a diminution of density as electrons on either side are accelerated and retarded respectively. They thus have minimum electron density when C has a maximum, and later show a maximum when electrons that started a half-cycle before or after them arrive in their neighbourhood, at the moment when there is minimum density in the neighbourhood of C . Hence, as the beam progresses, density maxima will occur in it twice in each period T of the applied field, in the neighbourhood of consecutive reference electrons in turn (spaced, in time, by $T/2$); so that a maximum occurs about ' C '-electrons at times $t_x, (T+t_x), \dots, (nT+t_x)$, and about ' A '-electrons at times $(t_x+T/2), \dots, [(n+\frac{1}{2})T+t_x]$.

The analogy will be evident with density fluctuations, or compressions and rarefactions, occurring in the passage of sound vibrations through a stationary medium. Indeed, it is helpful to view the beam as a uniform steady distribution of electrons upon which the alternating field impresses variations in density, the maxima and minima of which occur always at the same points in space: The modulated beam thus corresponds to the stationary wave produced by interaction of a direct and reflected sound-wave. It must be observed, however, that the change in speed of individual electrons is proportional to the root of the change in voltage which produces it, since it is the energy and not the velocity which is directly influenced by the field. The above considerations are only valid, therefore, if the maximum voltage fluctuation ΔV_m is small compared with the accelerating voltage V_0 applied to the beam; the change in velocity is then small compared with the initial velocity and has approximately the sinusoidal variation of the modulating field. Alternating voltages of saw-tooth or other wave-form will also produce bunching of the beam, but the Applegate diagram will differ from that given above.

Separation of Electrodes. The time t_x and distance s_x from the buncher at which the first pulse is formed may be found by the following simplified treatment, with reference to Fig. 150. The electron C continues with unaltered velocity v through the field-free drift space beyond the buncher, which it leaves at time $t = 0$. The electron D passes the buncher a quarter-cycle $T/4$ later, and acquires the maximum additional velocity Δv_m . On leaving the buncher it is a distance $v_0 T/4$ behind C and has a velocity Δv_m relative to it: it will overtake it after a time $v_0 T/4\Delta v_m$. Thus C and D are in closest proximity at time

$$t_{cd} = v_0 T/4\Delta v_m + \frac{1}{4}T = (v_0/\Delta v_m + 1)\frac{1}{4}T. \quad (X.1)$$

Similarly C overhauls electron B which passed the buncher a quarter-cycle in advance of it, suffering the maximum retardation of Δv_m . When C passes the buncher, B is a distance $(v_0 - \Delta v_m)T/4$ ahead of it and their relative velocity is Δv_m in favour of C . Hence these two electrons are in closest proximity at time

$$\begin{aligned} t_{bc} &= (v_0 - \Delta v_m)T/4\Delta v_m \\ &= \frac{1}{4}T(v_0/\Delta v_m - 1), \end{aligned} \quad (X.2)$$

so that C passes B a half-period earlier than it is in turn passed by D . A well-defined current pulse may therefore be said to exist during the time interval t_{bc} to t_{cd} . Since in any case the mutual repulsion of the electrons prevents very close approach, we may take the peak of current density to occur at the middle of this interval, at time

$$t_x = v_0 T/4\Delta v_m \approx \frac{1}{2}TV_0/\Delta V_m, \quad (X.3)$$

on substituting voltages for velocities. The distance from buncher to catcher should thus be

$$s_x = v_0 t_x = \frac{T}{4} \frac{v_0^2}{\Delta v_m} = \frac{\beta \lambda}{4} \frac{v_0}{\Delta v_m} = \frac{\beta \lambda}{2} \frac{V_0}{\Delta V_m}, \quad (X.4)$$

where λ is the generated wave-length and β the ratio of v_0 to the velocity of light.

As mentioned above, other current pulses will be formed at later intervals of $T/2$ in the motion of the beam, and the general expressions for the time and distance from the catcher of their occurrence will be

$$t_n = (\frac{1}{2}a - \frac{1}{2})Tv_0/\Delta v_m = \frac{1}{2}T(2a-1)v_0/\Delta v_m, \quad (\text{X.5})$$

$$\text{and} \quad s_n = \frac{1}{2}T(2a-1)v_0^2/\Delta v_m = \frac{\beta\lambda}{4}(2a-1)\frac{v_0}{\Delta v_m}, \quad (\text{X.6})$$

where a is any integer. The approximation made in arriving at the conditions for the initial pulse becomes less valid with increasing value of a and the pulses become less well defined, as is evident if the Applegate diagram be continued to greater distances. Hence it is desirable to set the collecting electrode at the minimum distance s_x from the buncher; under the usual conditions of operation this will be less than a wave-length. The position is not at all critical, however, since the pulse will be adequately sharp over a time equal to half the period of the alternating field, as shown by equations X.1 and X.2. By comparison with X.3, it will be seen that the latitude of adjustment allowed is equal to a fraction $2\Delta v_m/v_0$ of the total separation s_x , which under normal operating conditions will amount to several per cent. If this were not so, the conditions necessary for oscillation would be so stringent that the tube would be very difficult to operate. For a klystron of given dimensions s_x is fixed and the resonant frequency and period T can be varied only within a narrow range. It follows from X.3 and X.4 that any change in the accelerating voltage must be accompanied by a corresponding change in the modulating amplitude, a condition which is further discussed below.

It will be evident that a period of drift after the beam passes the bunching grids is essential to the operation of the system, allowing the velocity modulation to be converted into density modulation of the beam. An ordinary triode valve produces such pulses of density in a direct manner, by the influence of the varying grid potential on the thermionic current passing from cathode to anode. At very high frequencies the transit time of these electrons has an adverse effect on the maintenance of oscillations. The klystron type of valve overcomes these difficulties by separating the density modulation stage from the control of emission, in such a way that the power taken by the modulating grids is a minimum whereas that delivered to the collecting anode is very large.

71. The Klystron

If a collecting electrode be inserted in the path of a density modulated beam of the type just described, at the distance from the modulating electrodes given by equation X.4, then it will experience current impulses of maximum intensity at intervals of T , the period of the modulating voltage. The collector may be in the form of a transmission line of such

length, or cavity of such shape or size, that it resonates electromagnetically with impulses of this frequency, and oscillations will then rapidly develop in it. If the energy thus acquired from the beam be greater than that delivered to it by the modulator, the oscillations will be self-maintained

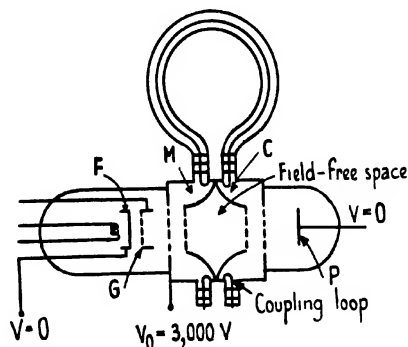


Fig. 151. Klystron oscillator: essential structure. (Varian and Varian.)

and the system will act as a generator of ultra-short waves. The speed of the electrons being of the order of 10^9 – 10^{10} cm. sec.⁻¹, the transit time between the electrodes will be less than 10^{-3} microsec., equivalent to a wave-length of less than 30 cm. A practicable oscillator of this form was devised by Varian and Varian and is known as the klystron; in later development it has proved capable of generating waves of a few cm. length only. On account of the inherent difficulties of operating a normal triode valve at frequencies of the order of 10^9 c.p.s., the klystron has found many uses, especially as it is more efficient than the magnetron in this range. Like the magnetron it turns to positive advantage the transit-time phenomena that limit the performance of conventional thermionic valves.

The essential structure of the klystron oscillator is shown in Fig. 151. In order to produce self-oscillations the collecting resonator must feed back energy to the modulator, and this is most efficiently effected if the latter is also in the form of a cavity resonator (or 'rhumbatron'). The cathode *F* and control grid *G* are therefore followed by a 'buncher' *M* in the form of a shallow double-walled cup, circular in cross-section, the base of which is covered by wire grids to admit the passage of the electron beam. The shape of the cavity is designed to give maximum efficiency of resonance (high *Q*-factor) and minimum input-impedance. The modulator is directly connected to the second resonator and both are several thousand volts positive in potential with respect to the cathode, so that the nearer surface of *M* acts as accelerating anode to the beam. In order to obtain high power output, two collecting electrodes are employed, the one to extract the high-frequency component and the other to absorb the direct-current energy of the beam. The former is in the form of a second resonator *C* identical in form with the modulator, but reversed in its attitude to the

beam and known as the 'catcher'. The other collector is a simple disk electrode P in the direct path of the beam and at cathode potential, so that a retarding field is applied between catcher and collector.

Oscillations will be set up in the catcher electrode if successive pulses in the density modulated beam pass through it at the resonant frequency of the cavity (or at one of its harmonics). The passage of the pulse across the annular gap in C induces a displacement current in the resonator, which quickly builds up in amplitude if the phase relations as well as the frequency are correct. Energy can only be delivered to C if electrons are retarded in passing through it, and the exchange will be a maximum if the centre of the current pulse passes the centre of the resonator gap at the peak of a negative half-cycle of its electric field. Ideally the amplitude of the peak should be just sufficient to bring to rest the slowest electrons of the pulse at the further grid of the gap; if the field were stronger than this, then these slow electrons would be accelerated in the reverse direction and would take energy from the field. The electrons that finally reach the collector plate P are therefore of low velocity and the heat dissipated is small.

It follows that for optimum working a current pulse must travel between buncher and collector in such a time that it arrives at the latter at the peak of a negative half-cycle. The reference electron at the centre of the pulse is that marked C in Fig. 150, passing the buncher as its field changes from negative to positive (at $t = 0$). The time of passage must therefore be $(n - \frac{1}{2})T$, where T is the period of oscillation and n is any integer. If the separation of the catcher and buncher is d and the velocity of the unaccelerated electron C is v_0 , then we have for the required time interval

$$d/v_0 = (n - \frac{1}{2})T \quad \text{and} \quad d = (n - \frac{1}{2}) \frac{v_0 \lambda}{c} = (n - \frac{1}{2}) \beta \lambda. \quad (\text{X.7})$$

Since d and T are fixed for a klystron of given dimensions, the catcher will collect energy only at certain values of the accelerating potential, according to the value of n . However, the resonator responds to a variety of harmonics, which are not usually simple multiples of the fundamental, and it can also be tuned within a certain range by varying the shape and size of the free cavity, so that condition X.7 is not in practice as rigid as at first appears.

On the other hand, X.4 gives the minimum distance s_x at which a defined pulse can be formed, in terms of the modulating conditions, and shows that in a given tube the magnitude of the accelerating potential at once fixes the modulating amplitude ΔV_m required for oscillation. From X.4 and X.7 we thus have the equality

$$\frac{V_0}{\Delta V_m} \frac{\beta \lambda}{2} = s_x = d = (n - \frac{1}{2}) \beta \lambda, \quad (\text{X.8})$$

or

$$\Delta V_m = \frac{2V_0}{4n - 1}. \quad (\text{X.9})$$

In effect, therefore, the modulating voltage is 'quantized' with these discrete values, although again in practice enough energy is delivered to the catcher over a range of velocities for this requirement not to be highly critical.

The energy acquired by resonator *C* is fed back to the modulator *M* by a quarter-wave concentric line, as shown in Fig. 151. So long as it is greater in quantity than the energy delivered to the beam by *M*, together with losses in the resonators and line, then a net increase in oscillatory energy occurs in each cycle and the oscillations will grow in amplitude, drawing energy from the high-voltage source employed to accelerate the beam. The high-frequency oscillations may be transferred to an external circuit by inserting a wire loop into the resonating cavity in a region where the magnetic flux density is high, so that high-frequency currents are induced in the loop. Detailed discussion of the theory of cavity resonators will be found in the works quoted at the end of the chapter. To start oscillations in the first place the resonators are tuned mechanically, if it is already known that the accelerating voltage is in the correct range for pulsations to pass from buncher to catcher in the time necessary for resonance. In other circumstances tuning may be a long process, since the correct voltage range is very small and therefore difficult to locate. It is then simpler to connect in series with the accelerating voltage supply an alternating source of amplitude sufficient to ensure that the total voltage will run into at least one range in which oscillation is possible. Oscillations are readily maintained once the correct conditions exist. The beam takes negligible energy from the buncher, over a whole cycle of the field, since as much is returned to it on the negative half-cycle by the retarded electrons as is expended in the positive half-cycle in accelerating other electrons; assuming, that is, that the entrant beam is of constant current density, so that equal numbers of electrons are retarded and accelerated. Thus a very small acquisition of energy by the catcher will allow oscillations to build up in the resonators.

The klystron is consequently a very efficient high-frequency oscillator, in comparison with a triode or magnetron. The power consumed is very small, owing to the symmetrical action of the buncher, and at the same time the output power can be made very large owing to the use of two collecting electrodes. The catcher removes most of the energy from the beam, without suffering heating from the impact of electrons such as occurs on the anode of a normal valve or magnetron. The direct current of the beam is absorbed by the second collector after retardation, and the heating effect here is correspondingly small. The theoretical efficiency of a klystron oscillator is given as 58 per cent.; a power output of 10 watts is easily obtained at 10 cm. wave-length. It may be modulated by applying signals to the control grid *G* before the buncher, but the variation of frequency with accelerating voltage makes such amplitude modulation difficult in practice. It is more usual to produce frequency modulation by applying a small modulating voltage in the cathode circuit.

passes the centre of the annular aperture of a cavity resonator *R* before impinging on the collecting plate *P*, which is at still higher positive potential. The resonator is in the form of a concentric quarter-wave line surrounding the outside of the evacuated tube. The beam is collimated by the action of a focusing electrode *F* and a longitudinal magnetic field due to the coil *M*. The tube is shown connected as an amplifier, the grid *G* being energized by a separate high-frequency source. It directly produces pulses of current density in the beam in the same way as in any triode oscillator, the accelerating potential between cathode *C* and anode *A* being constant. The passage of these pulses across the gap of the resonator sets up sympathetic oscillations in it, if the voltages are properly adjusted. Haeff employed an accelerating voltage of 6,000 V., and reported an output of 110 watts at an amplifying efficiency of 35 per cent., for a wave-length of 60-70 cm. If excitation of the grid is produced by feeding back energy from the resonator, self-oscillations may be maintained as in a klystron.

(b) The *Hahn-Metcalf* tube operates in a different manner, the beam being modulated and demodulated by grids in the form of short cylinders, between which a series of accelerating diaphragms and focusing electrodes keep the beam in the axis of the tube. The cathode *C* is surrounded by a first grid (1) at a fixed potential of 60 V. and an accelerating electrode (2) at 300 V., which ensure that the fluctuation of voltage on the modulating grid (3) does not affect the number of electrons emitted from the cathode (Fig. 154). The cylindrical electrodes 5, 7, 9, and 11, at 30 V. positive with respect to the cathode, form a series of electrostatic lenses, in conjunction with the diaphragms 4, 6, 8, 10, 12, and 14, all at 300 V. positive; the system is equivalent to a multi-element symmetrical lens. The beam is thus prevented from spreading during the period of drift, and then passes through the centre of the 'catching' anode 13 before reaching the collector 15. The grid 3 and anode 13 are both operated at a positive potential of from 10 to 30 V. An alternating potential imposed on the grid serves to modulate the velocity of the beam, so that current pulses are formed in the ensuing drift space, just as in the klystron, and finally deliver energy to the anode. If an external source drives the grid, the tube acts as an amplifier; it operates at frequencies of the order of 100 Mc.p.s. As with

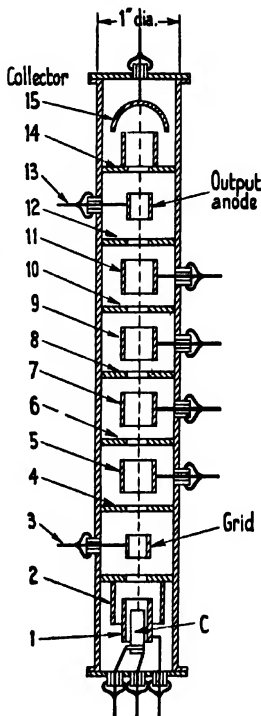


FIG. 154. Velocity modulation tube of Hahn and Metcalf.

other velocity-modulated tubes, however, the output may be fed back to the input grid and self-maintained oscillations then generated. Hahn and Metcalf describe the operation of such oscillators down to a wavelength of 4.8 cm.

The application of the principle of velocity-modulation is still in its early stages, as will be seen from the references. It has rapidly led to the production of new types of oscillators, operating at frequencies which are unattainable by the triode valve and yielding large output power. There can be little doubt that the incorporation of focusing devices, as in the tube last described, will allow of increased efficiency and flexibility of operation. Considerable improvements must have been made in the past few years, which have remained unpublished owing to the necessities of war. It is hoped that the brief description of fundamentals given in these pages will aid in the understanding of further information as it is released for publication.

FURTHER READING

- Brüche and Recknagel, *Z. Phys.* 101, 459, 1938.
Varian and Varian, *J. Appl. Phys.* 10, 321, 1939.
Webster, *J. Appl. Phys.* 10, 501, and 864, 1939.
Haeff, *Electronics*, 12, 30, 1939.
Haeff and Nergaard, *P.I.R.E.* 28, 126, 1940.
Hahn and Metcalf, *ibid.* 27, 106, 1939.
Brainerd, *Ultra-High-Frequency Techniques*, Chapter 10 (1942).
R.C.A. Staff, *Radio at Ultra-High-Frequencies* (1942).
Lamont, *Wave Guides* (1942).
Harvey, *High Frequency Thermionic Tubes* (1944).

APPENDIX

THE HAMILTONIAN METHOD IN ELECTRON OPTICS

A 1. *Introduction*

THE treatment of trajectories given in the text has been that of classical dynamics, based upon the corpuscular nature of an ion or electron. A more concise and exact discussion is provided by the theory of wave mechanics, as developed especially by Schrödinger, the foundations of which are the hypothesis of de Broglie and the experimental verification of it by electron diffraction (cf. section 58). The equations for the material waves associated with, or 'guiding', an electron lead to a straightforward expression (A.12, below) relating the 'refractive index' of a field to a characteristic function V when the wave-length is very small. This expression defines the surfaces of constant phase, or wave-fronts, the orthogonals of which give the permitted trajectories in the medium. It is referred to by continental authors as the 'eiconal equation', and corresponds to the Hamilton-Jacobi equation for the motion of a particle in a conservative field of force, that is, when the total energy of the particle is constant. There may then be deduced from it (see Picht, § 3, p. 25) an expression (A.3) which is essentially Fermat's principle of least time, familiar in optics.

From Fermat's principle Sir William Rowan Hamilton had, a century earlier, developed a complete theory of the passage of light through anisotropic heterogeneous media, wherein the refractive index not only varies continuously from point to point, but depends also on the instantaneous direction of a ray. As we have seen (section 36), the electromagnetic field behaves as a medium of this type towards charged particles, an expression for the meridional potential (= refractive index) being obtained (IV.51) which involves implicitly the direction as well as the position of a particle at any time, owing to the anisotropy introduced by the magnetic field. It is therefore legitimate to transfer Hamilton's method directly to the electron optical problem, the form of the expressions remaining unchanged and the special value of the refractive index being inserted. Consequently the rigorous wave-mechanical introduction is here omitted, the interested reader being referred to the short account in Picht's work and to a valuable application by Glaser of wave mechanics to image formation in the electron microscope.

Glaser has been responsible for developing electron optical theory on Hamiltonian lines, and his method will be followed here: Fermat's principle is taken as an axiom and the Hamilton-Jacobi equation deduced from it, instead of vice versa as in wave mechanics. It leads to an elegant treatment of aberrations, alternative to the trigonometrical method of

Scherzer, Chapter V (above), as well as allowing the deduction of the fundamental differential equation of the ray path. It will find increasing use in the discussion of electron optical problems, and therefore is treated in this Appendix. An excellent account of Hamilton's method in light optics is available in Synge's monograph in the Cambridge series of tracts. Although it does not deal with the general case of anisotropic heterogeneous media, which are not met with in practical optics, it will be found to give a helpful picture of the Hamiltonian standpoint, the last chapter being especially illuminating for its discussion of isotropic heterogeneous media. Unfortunately the notation employed on the Continent is not that of Hamilton; it differs also from author to author and even in different papers by the same author. Synge uses Hamilton's own notation and the same practice will be followed here, the variants of Glaser and Picht† being indicated when the symbols for the chief quantities first appear.

The principle of Fermat, on which Hamilton's optical work is based, states that the actual path along which light travels between two points has a stationary (maximum or minimum) optical length as compared with adjacent curves joining the two points. In his work on dynamics Hamilton used an analogous 'principle of least action', known after him, which states that a particle follows such a trajectory between two points that the 'action' has a stationary value in comparison with neighbouring paths joining the points. The action is defined as a time integral of energy, and Hamilton's action principle may be stated as

$$\int mv^2 dt = \int mv ds \rightarrow \text{maximum or minimum,}$$

or

$$\delta \int v ds = 0. \quad (\text{A.1})$$

His optical and dynamical theories being very closely related, this expression leads also to the Hamilton-Jacobi equation mentioned above and derived later (A.12). Using these concepts, Funk has developed an electrodynamical treatment of electron lenses on parallel lines to the wave theory of Glaser. Gray has used the same principles in simpler form in order to derive the form of field having minimum aberration. The method is equivalent to considering a flux of electrons, rather than a single particle, in an electromagnetic field and finding expressions for the possible paths between two given points. The details will not be reproduced here, the results being essentially the same as those obtained by Glaser, but Gray's work in particular will repay study, as he obtains approximate expressions for the conditions of minimum aberration in a form from which physical consequences may be immediately drawn.

† Picht in his *Einführung in die Theorie der Elektronenoptik* follows Glaser's early treatment, which differs in many respects from his later papers in the *Beiträge zur Elektronenoptik* and *Z. techn. Phys.* References to Glaser in this Appendix are to this later and more concise work.

A 2. The Point Characteristic, and the Fundamental Ray Equations

Fermat's principle of least time, as stated above, involves the optical length of the ray path. In its original form it referred to the time of passage, stating that this must have a stationary value along the actual path as compared with adjacent curves joining the same two points; but the time integral may be directly transformed into a path integral:

$$t = \int_{A_1}^{A_2} dt = \int_{A_1}^{A_2} ds/v = 1/c \int_{A_1}^{A_2} \mu ds \rightarrow \text{maximum or minimum.} \quad (\text{A.2})$$

The integral is taken along the true path of the ray, of which ds is an element, between two points A_1 and A_2 ; c is the velocity of light in free space and μ the refractive index of the medium for the (monochromatic) radiation used. The product (μds) is then defined as the optical length of a path element and the integral $\int_{A_1}^{A_2} \mu ds$ as the optical length of the complete path. Hence the optical length L is the path length in free space which would be travelled in the same time as is taken by an actual ray between two points in the given medium:

$$L = ct = \int_{A_1}^{A_2} \mu ds. \quad (\text{A.3})$$

By Fermat's principle the difference δL in the optical length of two neighbouring paths must be zero:

$$\delta L = \delta \int \mu ds = 0. \quad (\text{A.4})$$

These concepts and the criterion of the stationary path may be at once applied in electron optics to the discussion of the trajectories of charged particles in the electromagnetic field. In place of μ we write the refractive index N of the field, which depends on the position and direction of motion of the electron. Referring its position in rectangular coordinates x, y, z to an arbitrary origin O , then N will be a function of x, y, z , and of their first derivatives with respect to s , the path parameter:

$$N = N(x, y, z, dx/ds, dy/ds, dz/ds) = N(x, y, z, \alpha, \beta, \gamma),$$

if we write α, β , and γ for the direction cosines of the tangent to the path at this point. Let the path parameter have the value s_1 at A_1 and s_2 at A_2 , the two termini of the path; then we have for the path length

$$L = \int_{s_1}^{s_2} N_{(xyz)} (x'^2 + y'^2 + z'^2)^{1/2} ds,$$

using the prime notation for derivatives with respect to s ; or

$$\left. \begin{aligned} L &= \int_{s_1}^{s_2} w \, ds, \\ \text{if we put } w &= N_{(xyz)}(x'^2 + y'^2 + z'^2)^{\frac{1}{2}}. \end{aligned} \right\} \quad (\text{A.5})$$

(Glaser uses F , and Picht n , instead of the symbol w .)

Now consider a set of adjacent curves joining A_1 and A_2 , described in terms of the parameter s , which thus has the same terminal values s_1 and s_2 on each of them. The variation of path length on passing from one curve to its neighbour is

$$\begin{aligned} \delta L &= \int_{s_1}^{s_2} \delta w \, ds \\ &= \int_{s_1}^{s_2} \left(\sum \frac{\partial w}{\partial x'} \delta x' + \sum \frac{\partial w}{\partial x} \delta x \right) ds, \end{aligned}$$

if \sum indicates the sum of the three terms obtained by changing $x \rightarrow y \rightarrow z$, and $\delta x'$, $\delta y'$, $\delta z'$, δx , δy , δz are the infinitesimal increments obtained on passing from a point on one curve to the point on the next having the same value of s . Then, since

$$\delta x' = \frac{d}{ds} \delta x; \quad \delta y' = \frac{d}{ds} \delta y; \quad \delta z' = \frac{d}{ds} \delta z,$$

partial integration gives

$$\delta L = \left[\sum \frac{\partial w}{\partial x'} \delta x \right]_{s_1}^{s_2} - \int_{s_1}^{s_2} \sum \left(\frac{d}{ds} \frac{\partial w}{\partial x'} - \frac{\partial w}{\partial x} \right) \delta x \, ds. \quad (\text{A.6})$$

As the curves have common end-points, $\delta x = \delta y = \delta z = 0$ at these points and therefore the first term vanishes. By Fermat's principle the second term must vanish also along the true, or 'natural', ray from A_1 to A_2 , so that the differential equations which must be satisfied by the trajectory are:

$$\left. \begin{aligned} \frac{d}{ds} \frac{\partial w}{\partial x'} - \frac{\partial w}{\partial x} &= 0 \\ \frac{d}{ds} \frac{\partial w}{\partial y'} - \frac{\partial w}{\partial y} &= 0 \\ \frac{d}{ds} \frac{\partial w}{\partial z'} - \frac{\partial w}{\partial z} &= 0 \end{aligned} \right\} \quad (\text{A.7})$$

These are the Euler differential equations for the extremals of $\int w \, ds$. Noting that

$$\frac{d}{ds} \frac{\partial w}{\partial x'} - \frac{\partial w}{\partial x} = \frac{d}{ds} \left[\frac{N x'}{(x'^2 + y'^2 + z'^2)^{\frac{1}{2}}} \right] - \frac{\partial N}{\partial x} (x'^2 + y'^2 + z'^2)^{\frac{1}{2}} = \frac{d}{ds} \left(N \frac{dx}{ds} \right) - \frac{\partial N}{\partial x}, \quad (\text{A.8})$$

because $x'^2 + y'^2 + z'^2 = 1$ when the parameter s is the arc length of the path, and substituting for w from A.5, we have

$$\left. \begin{aligned} d/ds \left(N \frac{dx}{ds} \right) - \frac{\partial N}{\partial x} &= 0 \\ d/ds \left(N \frac{dy}{ds} \right) - \frac{\partial N}{\partial y} &= 0 \\ d/ds \left(N \frac{dz}{ds} \right) - \frac{\partial N}{\partial z} &= 0 \end{aligned} \right\} \quad (\text{A.9})$$

From this form of the Euler equations Picht (p. 43) proceeds to derive the differential equations of the electron path in the form otherwise arrived at above (IV.54, IV.55).

In the present discussion we are concerned with the variation in the path length L with the position of the object point A_1 and the corresponding image point A_2 . For this purpose it is convenient to replace L by the characteristic function V . If the coordinates of these two points are x_1, y_1, z_1 , and x_2, y_2, z_2 , respectively, then the characteristic function is defined, in the same way as L , as the optical length of the natural path $A_1 A_2$, and may be written:

$$V = V(x_1, y_1, z_1, x_2, y_2, z_2) = \int_{A_1}^{A_2} N ds. \quad (\text{A.10})$$

As it specifies the ray between two points it is known as the 'point characteristic' of Hamilton, or among continental writers as the 'point-eiconal'. Glaser uses the symbol E or W for it, and Picht the symbol \mathcal{E} .

An expression may now be obtained for the change produced in V by any infinitesimal displacement of A_1 to B_1 and A_2 to B_2 . If the terminal values of the parameter s remain the same for the new ray, the variation in the optical path length V is given by A.6 and, since the path $A_1 A_2$ is a natural ray, the integral must vanish. Then, as in A.8, we have

$$\partial w / \partial x' = Nx' = N\alpha = \sigma; \quad \partial w / \partial y' = Ny' = \tau; \quad \partial w / \partial z' = Nz' = \nu, \quad (\text{A.11})$$

where σ, τ , and ν , as defined by these equations, are styled the 'components' of the ray. The path difference is now given by the first term of A.6:

$$\delta V = [\sum \partial w / \partial x' \delta x]_{A_1}^{A_2} = \sum \sigma_2 \delta x_2 - \sum \sigma_1 \delta x_1,$$

where \sum as before indicates a sum of three terms involving x, y , and z identically. We then have for the partial derivatives

$$\begin{aligned} \frac{\partial V}{\partial x_1} &= -\sigma_1; & \frac{\partial V}{\partial y_1} &= -\tau_1; & \frac{\partial V}{\partial z_1} &= -\nu_1; \\ \frac{\partial V}{\partial x_2} &= \sigma_2; & \frac{\partial V}{\partial y_2} &= \tau_2; & \frac{\partial V}{\partial z_2} &= \nu_2. \end{aligned}$$

But since the direction cosines are related by the identities

$$\alpha_1^2 + \beta_1^2 + \gamma_1^2 = 1; \quad \alpha_2^2 + \beta_2^2 + \gamma_2^2 = 1,$$

we have also

$$\sigma_1^2 + \tau_1^2 + \nu_1^2 = N_1^2, \quad \text{and} \quad \sigma_2^2 + \tau_2^2 + \nu_2^2 = N_2^2,$$

and therefore

$$\left. \begin{aligned} \left(\frac{\partial V}{\partial x_1} \right)^2 + \left(\frac{\partial V}{\partial y_1} \right)^2 + \left(\frac{\partial V}{\partial z_1} \right)^2 &= N_1^2 \\ \left(\frac{\partial V}{\partial x_2} \right)^2 + \left(\frac{\partial V}{\partial y_2} \right)^2 + \left(\frac{\partial V}{\partial z_2} \right)^2 &= N_2^2 \end{aligned} \right\} \quad (\text{A.12})$$

These two partial differential equations are satisfied by the point characteristic function V . They are identical with the Hamilton-Jacobi equations obtained by a dynamical treatment for the motion of a particle in a conservative field of force, and form the basis of the alternative method of Funk and Gray already mentioned.

A 3. *The Mixed Eiconal: Equation of the Image*

The expressions in V given above define the surfaces of constant path length emanating from an object point A_1 , of which the orthogonal trajectories are the ray paths. The final equations (A.12) give the variation in path length with position of A_1 and A_2 , and therefore provide the possibility of determining the resulting aberration of the image point in terms of a difference in path length. However, this longitudinal aberration in the direction of the ray is not of immediate practical application. It is preferable to find the lateral aberration in the Gaussian (paraxial) image-plane, for which purpose a function must be used which is dependent on the direction rather than the coordinates of the ray at the object and image. Glaser's treatment proceeds from a different formulation of the first term in the path difference equation A.6, and leads to a definition of a second characteristic function, the 'mixed' characteristic (W), involving the coordinates of the object point and the direction cosines of the image point. Picht employs the third, or angle, characteristic T (W , in his notation), which is a function of direction cosines only.

It is convenient to refer all measurements to the z -axis, and to restrict discussion to fields of rotational symmetry. Expression A.5 for the optical path length of a ray then becomes:

$$V = \int_{A_1}^{A_2} N \sqrt{(x_1'^2 + y_1'^2 + 1)} dz = \int_{A_1}^{A_2} w dz, \quad (\text{A.13})$$

defining w in the same manner as before. Glaser proceeds as above to write for the variation in path length between two rays starting from the same object point A_1 (at x_1, y_1):

$$\delta V = \int_{A_1}^{A_2} \delta w dz = \int_{A_1}^{A_2} \left(\sum \frac{\partial w}{\partial x'} \delta x' + \sum \frac{\partial w}{\partial x} \delta x \right) dz,$$

where \sum now indicates the sum of *two* expressions involving x and y identically. On partial integration this gives the corresponding expression to A.6,

$$\delta V = \left[\sum \frac{\partial w}{\partial x'} \delta x \right]_{A_1}^{A_2} - \int_{A_1}^{A_2} \sum \left(\frac{d}{dz} \frac{\partial w}{\partial x'} - \frac{\partial w}{\partial x} \right) \delta x \, dz. \quad (\text{A.14})$$

The integral must vanish for a natural ray, so that if we fix attention on such a ray from A_1 to the image point A_2 (at x_2, y_2), then the first term gives the path difference between this and any other adjacent ray starting from A_1 :

$$\delta V = \left(\frac{\partial w}{\partial x'} \delta x + \frac{\partial w}{\partial y'} \delta y \right)_{A_2} - \left(\frac{\partial w}{\partial x'} \delta x + \frac{\partial w}{\partial y'} \delta y \right)_{A_1}. \quad (\text{A.15})$$

This may be expressed in vector form by defining a radial vector \mathbf{r} with components x and y , and a 'normal' vector \mathbf{n} with components σ and τ , where $\sigma (= \partial w / \partial x')$ and $\tau (= \partial w / \partial y')$ are the 'components' of the ray as defined by A.11. Hence

$$\delta V = (\mathbf{n}_2 \cdot \delta \mathbf{r}_2) - (\mathbf{n}_1 \cdot \delta \mathbf{r}_1), \quad (\text{A.16})$$

which is known as Hamilton's fundamental equation, since it contains implicitly all the laws of image formation. In an isotropic medium the normal vector will coincide with the tangent vector of the ray at A_1 (or A_2).

Rearrangement of the fundamental equation gives

$$\delta V = \delta(\mathbf{n}_2 \cdot \mathbf{r}_2) - (\mathbf{r}_2 \cdot \delta \mathbf{n}_2) - (\mathbf{n}_1 \cdot \delta \mathbf{r}_1),$$

$$\text{or} \quad \delta\{V - (\mathbf{n}_2 \cdot \mathbf{r}_2)\} = -(\mathbf{r}_2 \cdot \delta \mathbf{n}_2) - (\mathbf{n}_1 \cdot \delta \mathbf{r}_1),$$

which becomes

$$\begin{aligned} -\delta W &= (\mathbf{r}_2 \cdot \delta \mathbf{n}_2) + (\mathbf{n}_1 \cdot \delta \mathbf{r}_1), \\ &= (x_2 \delta \sigma_2 + y_2 \delta \tau_2) + (\sigma_1 \delta x_1 + \tau_1 \delta y_1), \end{aligned} \quad (\text{A.17})$$

$$\text{on writing} \quad W = V - (\mathbf{n}_2 \cdot \mathbf{r}_2) = \int_{A_1}^{A_2} w \, dz - (\mathbf{n}_2 \cdot \mathbf{r}_2). \quad (\text{A.18})$$

The function W as thus defined is known as the mixed characteristic (or eiconal), being a function of the coordinates of the object point (x_1, y_1) and the ray components of the image point (σ_2, τ_2). Glaser employs for it the symbol V , thus interchanging Hamilton's notation for the point and mixed functions; Picht does not use it, preferring the angle-characteristic (T , in Hamilton; π , V , or W in Picht). In light optics the latter is usually employed when the incident rays form a parallel beam, the mixed function when they are emitted from a point, as in the present discussion.

The meaning of W may be more obvious from a geometrical derivation. If $A_1 K A_2$ be a natural ray from A_1 to A_2 (Fig. 155), and O is the origin of coordinates, then \mathbf{r}_1 and \mathbf{r}_2 are the radial vectors defined above. Let the length and direction of the normal vectors \mathbf{n}_1 and \mathbf{n}_2 be as shown; they will not be tangent to the path, as the medium is supposed anisotropic. From

O let a perpendicular be dropped on to \mathbf{n}_2 to meet it in P_2 . Then the optical path length $V_{(A_1, P_2)}$ from A_1 to P_2 is given by

$$V_{(A_1, P_2)} = V_{(A_1, A_2)} + (N_2 \cdot A_2 P_2) = V_{(A_1, A_2)} - (\mathbf{n}_2 \cdot \mathbf{r}_2), \quad (\text{A.19})$$

since $A_2 P_2$ is the orthogonal projection of \mathbf{r}_2 on \mathbf{n}_2 . The right-hand side of the expression is now defined as the mixed characteristic $W_{(A_1, A_2)}$, in agreement with A.18. Hence the mixed function assigned to a given ray

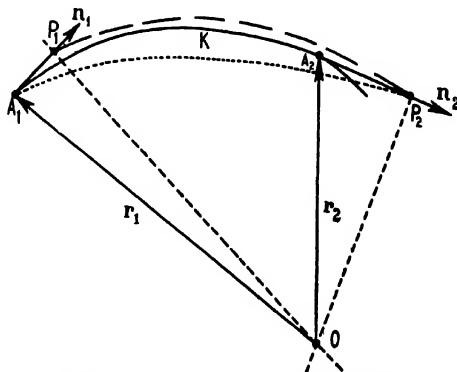


FIG. 155. Construction of mixed and angle characteristic functions.

path is identical with the point function of a path having the same starting-point but leading to the foot of the normal from the origin on to the final normal vector:

$$W_{(A_1, A_2)} = V_{(A_1, P_2)} = V_{(A_1, A_2)} - [A_2 P_2], \quad (\text{A.20})$$

where square brackets indicate an optical, not a geometrical, length. Similarly the angle (or 'normal') characteristic is defined as the path length between the feet (P_1, P_2) of the normals dropped from the origin on to the initial and final normal vectors respectively. When the media containing A_1 and A_2 are both homogeneous and isotropic, then the normal vectors lie in the initial and emergent rays respectively, and the different values of the three characteristic functions arise simply from selecting different points on the path of the ray between which to make measurements; this is the usual condition in light optics. In the electron optical case the three functions are equivalent to the optical length, or value of the point function, along three neighbouring rays having terminal points defined as above; the two imaginary rays will differ from the true path by amounts depending on the rapidity of variation of refractive index with position and direction in the object and image media.

The mixed characteristic allows the coordinates of the image point to be found. It follows from A.17 that

$$x_2 = -\partial W / \partial \sigma_2, \quad \text{and} \quad y_2 = -\partial W / \partial \tau_2, \quad (\text{A.21})$$

so that if W is known, as a function of x_1, y_1, σ_1 and τ_1 , the position of the

image point A_2 corresponding to A_1 can be found. It is to be noted that by its construction W is a function of the direction of the ray at the image not at the object, and it is therefore usually impossible to evaluate directly. Instead, we proceed by expanding W in ascending powers of the variables x_0, y_0, σ_2 , and τ_2 , taking only those terms determining the third order aberrations and finally replacing the angular variables by the physically more accessible coordinates of the ray in some arbitrary aperture plane. The expansion of W involves even powers only, the odd vanishing in the conditions of rotational symmetry assumed, so that

$$W = W_0 + W_2 + W_4 + \dots, \quad (\text{A.2})$$

where the subscripts indicate the order of the variables entering into the respective terms. Differentiation of this series with respect to σ_2 and τ_2 will then give the coordinates of the image point (by A.21) to any desired accuracy, according to the number of terms taken. By taking the first two terms only, $W_0 + W_2$, we obtain the image point for paraxial rays. By including also W_4 the image point is obtained with the same accuracy as by including the third-order term of the sine expansion in the trigonometrical method (cf. section 37). We denote this 'Seidel' image point by the coordinates $x_{2(4)}, y_{2(4)}$, and the first-order (or Gaussian) image point by the result $x_{2(2)}, y_{2(2)}$, indicating in brackets the highest term in A.22 which contributes to it. In the same way as in the text (p. 119) we then define the third-order aberration to be the displacement of the first point with respect to the second:

$$\Delta x_2 = x_{2(4)} - x_{2(2)}; \quad \Delta y_2 = y_{2(4)} - y_{2(2)}.$$

In each case the difference will be due to the derivative of the fourth-order term in the expansion of W , and the aberrations are given simply by

$$\Delta x_2 = -\partial W_4 / \partial \sigma_2; \quad \Delta y_2 = -\partial W_4 / \partial \tau_2. \quad (\text{A.23})$$

The term W_4 , however, has still to be found by expansion of the integral for W (A.18), and ambiguity remains as to the path over which the integral is to be taken: that of the paraxial ray, or of some closer approximation to the true path. It proves to be sufficiently accurate for third-order theory to take the former. By a theorem which will not be demonstrated here (see Synge, p. 66) a small change in the path length affects W (and V) only in proportion to its square, so that restriction to the Gaussian conditions, dropping terms of third and higher orders in z and y , is equivalent to neglecting terms of sixth and higher orders in W . Hence the value obtained for W_4 is not affected by employing in the expansion the integral of w taken along the paraxial path; linear instead of cubic equations are thus obtained for the aberration, and the solution greatly simplified.

A 4. *Seidel's Characteristic: the Total Aberration Equation*

An attempt to evaluate the aberration from A.23 is impracticable, since the components of the ray at the image point are involved. It is preferable now to introduce the coordinates (x_B, y_B) of the ray in some convenient

aperture plane (at z_B) in place of the components σ_2 and τ_2 . Owing to the linearity of the solutions for the paraxial ray equation (cf. section 36), we may replace W_4 directly by S_4 , a function of x_1, y_1, x_B and y_B , and write

$$\Delta x_2 = -\partial S_4 / \partial x_B; \quad \Delta y_2 = -\partial S_4 / \partial y_B, \quad (\text{A.24})$$

from A.23. By analogy with light optics the quantity S_4 is known as Seidel's characteristic function or eiconal; the same symbol is used here by both Glaser and Picht.

The Seidel function depends on x_1, y_1, x_B and y_B , but not all values of these variables are admissible. In view of the rotational symmetry of the system the value of S (and that of W) is determined by the radial distance of the ray from the axis and not by independently varying x and y coordinates. The four variables can therefore only appear in certain combinations: if r_1 is the radial coordinate of the object point, and r_B that of the ray in the aperture, then the possible combinations are

$$R = \mathbf{r}_1^2 = x_1^2 + y_1^2; \quad \rho = \mathbf{r}_B^2 = x_B^2 + y_B^2;$$

$$\kappa = \mathbf{r}_1 \cdot \mathbf{r}_B = x_1 x_B + y_1 y_B; \quad \xi = \mathbf{r}_1 \times \mathbf{r}_B = x_1 y_B - y_1 x_B;$$

R, ρ, κ , and ξ are convenient symbols for use in the subsequent expansion which gives S_4 . It will be seen that κ and ξ are the scalar and vector product respectively of the radius vectors \mathbf{r}_1 and \mathbf{r}_B , as indicated in the usual notation. As only fourth-order products of the variables enter into S_4 , it must have the form

$$S_4 = -\frac{1}{4}AR^2 - \frac{1}{4}B\rho^2 - C\kappa^2 - \frac{1}{2}DR\rho + ER\kappa + F\rho\kappa + GR\xi - H\kappa\xi + J\rho\xi, \quad (\text{A.25})$$

where A, B , etc., are coefficients depending on the refractive index and therefore, in the electron optical case, on the form of the field; a term in ξ^2 does not appear, being already included in terms C and D . The negative signs and numerical factors are inserted in reference to the expressions for the coefficients derived below. Partial differentiation then gives the aberrations in the x - and y -directions by A.24:

$$\begin{aligned} \Delta x_2 &= -\frac{\partial S_4}{\partial x_B} = B\rho x_B + Dx_1^2 x_B + 2Cx_1^2 x_B - Ex_1^3 - F\rho x_1 - 2Fx_1 x_B^2 + \\ &\quad + Hx_1^2 y_B - 2Jx_1 x_B y_B; \\ \Delta y_2 &= -\frac{\partial S_4}{\partial y_B} = B\rho y_B + Dx_1^2 y_B - 2Fx_1 x_B y_B - Gx_1^3 + Hx_1^2 x_B - J\rho x_1 - \\ &\quad - 2Jx_1 y_B^2; \end{aligned} \quad (\text{A.26})$$

if we follow Scherzer in setting $y_1 = 0$, so that the object point is on the x -axis; we may then set $x_1 = r_a$, as in section 38. Then the total aberration, or the displacement of the third-order image point with respect to the paraxial position, is given by

$$\begin{aligned} \Delta x_2 + i\Delta y_2 &= B(x_B^2 + y_B^2)(x_B + iy_B) + (C+D)r_a^2(x_B + iy_B) + \\ &\quad + (C+iH)r_a^2(x_B - iy_B) - (E+iG)r_a^3 - \\ &\quad - 2(F+iJ)r_a(x_B^2 + y_B^2) - (F-iJ)r_a(x_B + iy_B)^2; \end{aligned}$$

or, writing $x_B + iy_B = u_B$, $x_B - iy_B = \bar{u}_B$, where $u_B \bar{u}_B = |\mathbf{r}_B|^2 = r_B^2$, the square of the radial distance of the ray in the chosen aperture, and $r_a = u_a = \bar{u}_a = |\mathbf{r}_1|$,

$$\Delta x_2 + i\Delta y_2 = \Delta u_2 = Bu_B^2 \bar{u}_B - 2(F + iJ)r_a u_B \bar{u}_B - (F - iJ)r_a u_B^2 + \\ + (C + iH)r_a^2 \bar{u}_B + (C + D)r_a^2 u_B - (E + iG)r_a^3. \quad (\text{A.27})$$

Equation A.27 is formally identical with that given by Scherzer's method (V.20), and the two sets of coefficients used to specify the individual aberrations may be compared directly. Scherzer's terms contain the magnification M , which is implicitly unity in the above treatment; they also involve the initial potential V_0 and special solution $r_{\alpha B}$ in a slightly different manner. Putting Scherzer's symbols first, the several terms refer as follows to the eight aberrations discussed in Chapter V:

$$\left. \begin{array}{l} \text{distortion:} \\ \text{anisotropic distortion:} \end{array} \right\} \alpha \equiv -\frac{MV_0}{r_{\alpha B}^2}(E + iG),$$

$$\text{curvature of field:} \quad \beta \equiv \frac{M\sqrt{V_0}}{r_{\alpha B}}(C + D),$$

$$\left. \begin{array}{l} \text{astigmatism:} \\ \text{anisotropic astigmatism:} \end{array} \right\} \gamma \equiv \frac{M\sqrt{V_0}}{r_{\alpha B}}(C + iH),$$

$$\left. \begin{array}{l} \text{coma:} \\ \text{anisotropic coma:} \end{array} \right\} \left\{ \begin{array}{l} \delta \equiv -M(F - iJ), \\ \epsilon \equiv -2M(F + iJ), \end{array} \right.$$

$$\text{spherical aberration:} \quad \zeta \equiv \frac{Mr_{\alpha B}}{\sqrt{V_0}}B.$$

Picht employs the same coefficients as those of this section, apart from the use of $-I$ in place of H , and H instead of J ; Glaser uses G_{00} , G_{01} and G_{11} for G , H , and J .

In order to obtain detailed expressions for these coefficients it is necessary to insert the equivalent refractive index of the electromagnetic field in equation (A.18) for the characteristic function W , and then to carry out the complete expansion required in the above derivation (cf. A.23). It has been shown in Chapter IV (section 36) that the motion of an electron in combined fields of rotational symmetry can be discussed in terms of a meridional potential Q , given by IV.51:

$$Q = \frac{e}{m}V - \frac{1}{2r^2} \left(\frac{e}{m}rA + C \right)^2,$$

where A is the magnetic vector potential and V the electrostatic potential at any point (r, z) , and C is a constant determined by the initial rotation of the electron about the axis. When the origin is on the z -axis, as here assumed, C vanishes and we have

$$Q = \frac{e}{m}V - \frac{1}{2} \frac{e^2 A^2}{m^2} \\ = \frac{1}{2} v^2 - \frac{1}{2} \frac{e^2 A^2}{m^2}, \quad (\text{A.28})$$

on substituting velocity v for potential, from the energy relation. Since the equivalent refractive index of the field is proportional to the root of the potential, we may then write

$$N_M = \sqrt{\left(v^2 - \frac{e^2 A^2}{m^2}\right)}.$$

According to the conditions of the discussion of section 36, this expression gives the index of refraction in a meridian plane rotating around the axis with the same angular velocity as the electron. More generally Glaser shows that the refractive index N_{rz} at any point (r, z) in the field is given by a similar expression

$$\begin{aligned} N_{(rz)} &= \sqrt{V_{rz}} - \frac{e}{m} \mathbf{A} \cos \phi \\ &= \sqrt{V_{rz}} - \frac{e}{m} \left(\mathbf{A} \cdot \frac{d\mathbf{r}}{ds} \right), \end{aligned} \quad (\text{A.29})$$

where ϕ is the angle between the magnetic vector potential and the direction of the path at this point, which is otherwise given by $d\mathbf{r}/ds$. Expressed thus, it is more obvious that the value of the refractive index is a function of direction as well as position, so that the Hamiltonian theory of anisotropic heterogeneous fields may be directly applied to the electromagnetic case. In terms of the present discussion we then have for A.18, neglecting $(\mathbf{n}_2 \cdot \mathbf{r}_2)$ in the first instance:

$$W = \int_{A_1}^{A_2} w \, dz = \int_{A_1}^{A_2} N \, ds = \int_{A_1}^{A_2} \left(\sqrt{V_{(rz)}} - \frac{e}{m} (\mathbf{A} \cdot \mathbf{s}) \right) ds,$$

where \mathbf{s} is the unit vector in the ray direction.

Going over to cylindrical coordinates (z, r, ψ) , we may write $\mathbf{A} = A_{(rz)}$ in the rotationally symmetrical field, and thus

$$\begin{aligned} W &= \int_{A_1}^{A_2} \left[\sqrt{\{V_{(xyz)}\sqrt{(x'^2 + y'^2 + 1)}\}} - \frac{e}{m} A_{(rz)} r \frac{d\psi}{dz} \right] dz \\ &= \int_{A_1}^{A_2} \left[\sqrt{\{V_{(rz)}\sqrt{(r'^2 + (r\psi')^2 + 1)}\}} - \frac{e}{m} A_{(rz)} r \psi' \right] dz. \end{aligned} \quad (\text{A.30})$$

Reference should be made to Glaser's work for details of the expansion of this expression, giving the fourth-power terms constituting the required function S_4 and, by partial differentiation, the several aberration terms; the end result only will be quoted here. In the process $V_{(rz)}$ and $A_{(rz)}$ are replaced by the first few terms of their representation as series in the axial values $V_{(0z)}$ and $H_{(0z)}$ respectively (IV.43 and IV.44).

As proved in section 36 (p. 109) any ray path satisfying the general differential ray equation (IV.54) may be expressed in terms of two special solutions. Glaser chooses for the latter the solutions $g_{(z)}$ and $h_{(z)}$, specified

by the following values at the origin (z_1) and in the aperture (z_B), and corresponding to Scherzer's r_y and r_a respectively (cf. p. 120):

$$\begin{aligned} g(z_1) &= 1; & g(z_B) &= 0; \\ h(z_1) &= 0; & h(z_B) &= 1. \end{aligned}$$

These two solutions are used to specify the paraxial path $A_1 A_2$ which forms the basis of the above discussion, the aberration being expressed as the extent of the departure of the third-order path from the paraxial path at the Gaussian image plane. Glaser thus finds the values of the aberration coefficients entering into A.28 to be

$$\begin{aligned} B &= \frac{1}{h_2' \sqrt{V_2}} \int_{z_1}^{z_2} (Lh^4 + 2Mh^2 h'^2 + Nh'^4) dz; \\ C &= \frac{1}{h_2' \sqrt{V_2}} \int_{z_1}^{z_2} (Lg^2 h^2 + 2Mgg'h'h' + Ng'^2 h'^2 - h_1'^2 V_1 K) dz; \\ D &= \frac{1}{h_2' \sqrt{V_2}} \int_{z_1}^{z_2} (Lg^2 h^2 + M(g^2 h'^2 + g'^2 h^2) + Ng'^2 h'^2 + 2h_1'^2 V_1 K) dz; \\ E &= -\frac{1}{h_2' \sqrt{V_2}} \int_{z_1}^{z_2} (Lg^3 h + Mgg'(gh)' + Ng'^3 h') dz; \\ F &= -\frac{1}{h_2' \sqrt{V_2}} \int_{z_1}^{z_2} (Lgh^3 + Mhh'(gh)' + Ng'h'^3) dz; \\ G &= \frac{h_1'}{h_2'} \sqrt{\frac{V_1}{V_2}} \int_{z_1}^{z_2} (Pg^2 + Qg'^2) dz; \\ H &= 2 \frac{h_1'}{h_2'} \sqrt{\frac{V_1}{V_2}} \int_{z_1}^{z_2} (Pgh + Qg'h') dz; \\ J &= \frac{h_1'}{h_2'} \sqrt{\frac{V_1}{V_2}} \int_{z_1}^{z_2} (Ph^2 + Qh'^2) dz. \end{aligned}$$

The first coefficient A (cf. A.25), which contains neither x_B nor y_B and hence disappears in the partial differentiation which yields the aberrations (A.26), has the form

$$A = \frac{1}{h_2' \sqrt{V_2}} \int_{z_1}^{z_2} (g^4 L + g'^4 N + 2g^2 g'^2 M) dz.$$

In these expressions the functions K , L , M , N , P , and Q , which derive from the refractive index (A.30), contain only the axial values of the

electrostatic and magnetic field, apart from electronic constants. They have the form

$$\begin{aligned}
 K &= \frac{e}{16mc^2} \frac{H_{0z}^2}{\sqrt{V_{0z}^3}}, \\
 E &= \frac{1}{32\sqrt{V_{0z}}} \left[\frac{1}{V_{0z}} \left(V_{0z}'' + \frac{e}{2mc^2} H_{0z}^2 \right)^2 - \frac{2e}{mc^2} H_{0z} H_{0z}'' - V_{0z}^{iv} \right]; \\
 M &= \frac{1}{8\sqrt{V_{0z}}} \left[V_{0z}'' + \frac{e}{2mc^2} H_{0z}^2 \right]; \\
 N &= \frac{1}{2} \sqrt{V_{0z}}; \\
 P &= \frac{1}{16} \sqrt{\left(\frac{e}{2mc^2 V_{0z}} \right) \left[\frac{H_{0z}}{\sqrt{V_{0z}}} \left(V_{0z}'' + \frac{e}{2mc^2} H_{0z}^2 \right) - H_{0z}'' \right]}; \\
 Q &= \frac{H}{4} \sqrt{\left(\frac{e}{2mc^2 V_{0z}} \right)}.
 \end{aligned}$$

In comparing these expressions with Picht's treatment it must be remembered that Glaser does not confine consideration to the electron and hence inserts no negative sign for e , and that K differs by a factor of 4 in the two treatments. Further, the function Q employed in this sense is to be distinguished from the meridional potential in A.28, for which the same symbol is used.

The discussion of individual aberrations continues in terms of the expressions given above on entirely the same lines as in Chapter V, and leading necessarily to equivalent values for the parameters of each error. Glaser, for instance, proceeds to obtain concise expressions for the spherical aberration of the magnetic field acting alone (cf. V.23). Gray resorts to approximation in order to obtain convenient expressions for the minimum conditions for certain aberrations. The final relations are effectively the same whether derived by the trigonometrical method of Chapter V or by the Hamiltonian treatment. The advantage of the latter lies in its mathematical elegance; it has not so far found application in the practical problems of ray-tracing.

FURTHER READING

- Glaser, *Z. Phys.* **83**, 104, 1933, and **97**, 177, 1935; **121**, 647, 1943 (electron microscope).
 Glaser, *Beiträge zur Elektronenoptik*, p. 24 (1937).
 Picht, *Einführung in die Theorie der Elektronenoptik* (1939).
 Sygne, *Geometrical Optics: an Introduction to Hamilton's Method* (Cambridge Tracts in Mathematics and Mathematical Physics, No. 37) (1937).
 Funk, *Mon. Math. Phys.* **43**, 305, 1936.
 Gray, *Bell Syst. Techn. J.* **18**, 1, 1939.
 Gabor, *P.I.R.E.* **33**, 792, 1945.

GENERAL BIBLIOGRAPHY

1. Brüche and Scherzer, *Geometrische Elektronenoptik* (1934).
The first treatise on the subject.
2. *Beiträge zur Elektronenoptik* (edited by Busch and Brüche) (1937).
A symposium on the theory and applications of electron optics, especially valuable for a concise presentation of aberration theory by both the electron-mechanical and Hamiltonian methods; also for an article on the experimental demonstration of aberrations, by Diels and Wendt.
3. Klemperer, *Electron Optics* (1939).
An excellent monograph in the Cambridge Physical Tracts series, aimed at research workers.
4. Myers, *Electron Optics* (1939).
A comprehensive treatment of the subject, especially valuable for the historical discussion of applications and for an exhaustive bibliography of original work.
5. Maloff and Epstein, *Electron Optics in Television* (1939).
Contains a good exposition of the fundamental principles and a detailed discussion of their practical application in the design of cathode-ray tubes.
6. Picht, *Elektronenoptik* (1939).
A concise presentation of the essential theory, including a chapter on the Hamiltonian treatment of electron optics; difficult reading. Reprinted in U.S.A. (1944) under licence from the Custodian of Enemy Property.
7. Zworykin and Morton, *Television* (1940).
A very full account of television apparatus and technique, including excellent introductory chapters on fundamental principles.
8. Millman and Seely, *Electronics* (1941).
A good general account of the principles and practice of electronic devices used in radio.
9. M.I.T. Staff, *Applied Electronics* (1943).
Similar to Millman and Seely, dealing with the main branches of radio technique, but is valuable also for chapters on electrons in metals and the conduction of electrons in gases.

TABLE III

Electronic Constants

Velocity of light $c = 2.9978 \times 10^{10}$ cm. sec.⁻¹

Electronic charge $e = 4.8025 \times 10^{-10}$ e.s.u.

$$= 1.6020 \times 10^{-20} \text{ e.m.u.}$$

Specific electronic charge $e/m_0 = 1.7592 \times 10^7$ e.m.u. gm.⁻¹

Rest mass of electron $m_0 = 9.1066 \times 10^{-28}$ gm.

Equivalent wave-length of electron $\lambda = \sqrt{(150/V)} \times 10^{-8}$ cm. (V in volts).

Energy of 1 electron-volt (eV.) $= 1.6020 \times 10^{-12}$ erg

$$\equiv 1.169 \times 10^4 \text{ degrees K.}$$

Energy of photon in volts $V = h\nu/e = \frac{1.2395}{\lambda} \times 10^4$ V. (λ in Angstroms).

Specific charge of proton $e/M_p = 9.5788 \times 10^3$ e.m.u. gm.⁻¹

Specific charge of alpha-particle $e/M_\alpha = 4.8223 \times 10^3$ e.m.u. gm.⁻¹

Planck's constant $h = 6.624 \times 10^{-27}$ erg sec.

Avogadro's number $N = 6.0228 \times 10^{23}$ mole⁻¹.

Boltzmann's constant $k = 1.3805 \times 10^{-16}$ erg deg.⁻¹

(Values taken from R. T. Birge, *Reports on Progress in Physics*, vol. viii, p. 90, 1941.)

Table IV

Velocity, Momentum, and Wave-length of Electrons

V = accelerating potential. v = velocity of electron. $\beta = v/c$. Hr = electronic momentum, if the electron describes a circular path of radius r in a uniform magnetic field of strength H . (This table has been calculated from the values of the electronic constants given above, and using the relativistic equations.)

V (volts)	$\frac{1}{\sqrt{1-\beta^2}}$	$v \times 10^{-10}$ (cm. sec. ⁻¹)	Hr (oersted cm.)	λ (Angstrom units)
1	1.0000020	0.005932	3.372	12.26
10	1.0000196	0.01876	10.66	3.878
20	1.0000391	0.02653	15.08	2.742
30	1.0000587	0.03250	18.47	2.239
40	1.0000783	0.03752	21.33	1.938
50	1.0000978	0.04195	23.85	1.735
60	1.0001174	0.04595	26.13	1.583
70	1.0001370	0.04963	28.22	1.465
80	1.0001565	0.05306	30.16	1.371
90	1.0001761	0.05628	31.99	1.293
100	1.0001957	0.05932	33.73	1.226
200	1.000391	0.08388	47.71	0.8668
300	1.000587	0.1028	58.46	0.7075
400	1.000783	0.1186	67.48	0.6127
500	1.000978	0.1326	75.47	0.5479
600	1.001174	0.1453	82.70	0.5000
700	1.001370	0.1569	89.35	0.4629
800	1.001565	0.1678	95.55	0.4328
900	1.001761	0.1779	101.3	0.4080
1,000	1.001957	0.1873	106.7	0.3877
2,000	1.00391	0.2645	150.9	0.2740
3,000	1.00587	0.3234	184.9	0.2237
4,000	1.00783	0.3730	213.7	0.1935
5,000	1.00978	0.4166	239.1	0.1730
6,000	1.01174	0.4555	262.0	0.1578
7,000	1.01370	0.4911	283.0	0.1461
8,000	1.01565	0.5245	302.8	0.1366
9,000	1.01761	0.5554	321.3	0.1287
10,000	1.01957	0.5846	338.9	0.1220
2.10 ⁴	1.0391	0.8151	481.6	0.08588
3	1.0587	0.9846	592.7	0.06978
4	1.0783	1.1216	687.5	0.06014
5	1.0978	1.237	772.2	0.05355
6	1.1174	1.338	850.1	0.04865
7	1.1370	1.427	922.0	0.04485
8	1.1565	1.506	990.3	0.04177
9	1.1761	1.578	1055	0.03919
10	1.1957	1.644	1117	0.03702
2.10 ⁵	1.3915	2.084	1649	0.02508
3	1.5873	2.329	2101	0.01968
4	1.7830	2.482	2515	0.01644
5	1.9786	2.588	2911	0.01421
6	2.174	2.663	3292	0.01256
7	2.370	2.718	3662	0.01129
8	2.565	2.761	4027	0.01027
9	2.761	2.795	4388	0.009423
10	2.957	2.822	4742	0.008720
2.10 ⁶	4.915	2.935	8200	0.005044
4	8.830	2.979	14950	0.002765
6	12.74	2.989	21660	0.001909
8	16.65	2.993	28340	0.001459
10	20.57	2.995	35030	0.001181

LIST OF SYMBOLS

THE dot and prime system is used throughout the book to denote differentiation. One or more dots ($\dot{}$, $\ddot{}$) indicate the first or higher derivative with respect to time (e.g., $\dot{z} = dz/dt$; $\ddot{z} = d^2z/dt^2$). One or more primes (\prime , $\prime\prime$) similarly represent derivatives with respect to axial distance ($V' = dV/dz$; $V'' = d^2V/dz^2$). Derivatives of higher order than the third are usually indicated by an index in small Roman numerals (e.g., $d^4V/dz^4 = V^{IV}$).

Vector quantities, when it is necessary to distinguish them, are printed in heavy type. The scalar product of two vectors (**A** and **B**) is then indicated by a dot (**A**·**B**) and their vector product by a cross (**A**×**B**).

<i>Symbol</i>	<i>Significance</i>
A	= Angstrom unit (10^{-8} cm.)
A	= magnetic vector potential.
A_r, A_z, A_ϕ	= components of magnetic vector potential in cylindrical coordinates.
<i>A</i>	= aberration coefficient.
a, A, A_j	= constants.
<i>a</i>	= form constant of magnetic field.
<i>a</i>	= linear acceleration.
a_r, a_z, a_ϕ	= components of acceleration in cylindrical coordinates.
<i>B</i>	= aberration coefficient.
b, b_l, B, B_1	= constants.
$B_{(k)}$	= coefficient of Bessel's function.
$C, C_{12}, C_{1z}, C_{22}$	= constant (usually of integration).
<i>C</i>	= aberration coefficient.
C_c	= parameter of chromatic aberration.
C_s	= parameter of spherical aberration.
<i>c</i>	= velocity of light.
c_1, c_2, c_a, c_β	= constants in ray equation.
<i>D</i>	= aberration coefficient.
<i>D</i>	= diameter of aperture.
d, d_s, d_T	= linear separation.
<i>d</i>	= form constant of magnetic field.
<i>d</i>	= crystal lattice spacing.
d_0	= linear resolving power.
E, E_a, E	= electric field strength.
E_x, E_y, E_z	= electric field strength in rectangular coordinates.
E_r	= radial electric field strength.
<i>E</i>	= charge of an ion.
<i>E</i>	= aberration coefficient.
e_s, e_m, e	= electronic charge.
<i>e</i>	= base of natural logarithms (= 2.71828...).
F	= force (usually, vector sum of electric and magnetic forces).
F_r, F_z, F_ϕ	= components of force, in cylindrical coordinates.
<i>F</i>	= potential function.
<i>F</i>	= aberration coefficient.
F_c	= coil factor.
f, f_1, f_2, F	= focal length of lens or mirror.
F_c	= centripetal force.

<i>Symbol</i>	<i>Significance</i>
F_m	= magnetic force.
f	= frequency of oscillation.
f_0	= resonant frequency.
G	= aberration coefficient
G	= potential function.
g, g_1, g_2, g_z	= special solution of ray path equation.
H, H_a, H_c	= magnetic field strength
H_r, H_z, H_ϕ	= magnetic field strength in cylindrical coordinates.
H_x, H_y, H_z	= magnetic field strength in rectangular coordinates.
H_0	= maximum field strength on axis of magnetic lens.
H_{rz}	= magnetic field strength at any point (r, z) .
H_{az}	= magnetic field strength at any point (z) on the axis.
r, H_{rz}	= radial magnetic field at a point (r, z) .
z, H_{rz}	= axial magnetic field at a point (r, z) .
H	= aberration coefficient.
h	= Planck's constant.
h, h_1, h_2, h_z	= special solution of ray path equation.
I, I_0, I_1	= current.
i, i_1, i_2, i_0	= current density.
i	= direction of measurement displaced through one right angle with respect to reference direction.
i	= angle of incidence.
J	= aberration coefficient.
J_0	= Bessel's function of zero order.
K	= Kelvin scale of temperature.
K	= field function (in Hamiltonian theory).
k	= constant.
k	= lens power parameter.
k	= Boltzmann's constant.
L	= optical path length.
L	= linear separation.
L	= field function (in Hamiltonian theory).
l	= length of deflecting plate.
M	= mass of an ion.
M	= field function (in Hamiltonian theory).
M, M_0, M_1	= magnification.
m	= electronic mass.
m_0	= electronic rest mass.
N_m	= refractive index in rotating meridional plane.
N, N_1, N_2, \dots, N_n	= refractive index of successive media.
N_{rz}, N_{xyz}	= refractive index at point (r, z) or (x, y, z) .
$N(W)$	= distribution of kinetic energies, among electrons.
N	= field function (in Hamiltonian theory).
n	= integer.
n	= normal vector.
P	= field function.
$P(W)$	= probability that a state of given energy W will be occupied.
p	= object distance.
p, p_1, p_0	= momentum.
Q	= meridional potential of electromagnetic field.
Q	= field function (in Hamiltonian theory).

<i>Symbol</i>	<i>Significance</i>
q	= image distance.
R	= ray path parameter.
R, R_n, R_0	= potential function.
r, R, R_z	= radius of curvature.
r	= angle of refraction.
$\mathbf{r}, \mathbf{r}_1, \mathbf{r}_B$	= radius vector.
$r, r_a, r_b, r_c, r_L, r_m, r_n, r_o, r_p, r_q$	} = radial distance from axis.
r_a, r_β, r_γ	
S	= special solution of ray path equation.
S	= Seidel's characteristic function (eiconal).
s, s_x, S	= linear separation.
T	= field function.
T	= Hamiltonian characteristic function (angle eiconal).
T	= temperature on Kelvin scale.
t, t_x	= time.
U	= equivalent electrostatic potential of magnetic lens.
u	= object distance.
u, u_B, \bar{u}_B	= complex number ($= x \pm iy$).
V	= Hamiltonian characteristic function (point eiconal).
V, V_i	= potential at a point.
V_{xyz}	= potential at any point (x, y, z) .
V_{rz}	= potential at any point (r, z) .
V_z	= potential at a point on the z -axis.
V_k	= potential difference in kilovolts.
V_m	= peak voltage of an alternating supply.
$V_0, V_1, V_2, \dots, V_n$	= potential of successive electrodes.
V_0	= work function of surface at 0°K .
V_e	= emission voltage of thermionic electrons.
v	= image distance.
v, v_r	= linear velocity.
v_x, v_y, v_z	= components of velocity, in rectangular coordinates.
W	= Hamiltonian characteristic function (mixed eiconal).
W, W_i, W_a, W_0, W_x	= energy of electron.
w	= optical path function.
x, x_i, x_L, x_0	= distance on x -axis.
y, y_i, y_L, y_0	= distance on y -axis.
Y	= linear deflexion, in cathode ray tube.
Y	= Neumann function.
z, z_i, z_L, z_0	= distance on z -axis.
z_f	= distance from centre of lens to principal focus.
z_p	= distance from centre of lens to principal plane.
α	= an angle.
$\alpha, \alpha_s, \alpha_o, \alpha_D$	= angular width of electron beam.
α	= angular resolving power.
α	= aberration coefficient (distortion).
β	= aberration coefficient (curvature of field).
β	= ratio of particle velocity to velocity of light (v/c).
γ	= aberration coefficient (astigmatism).
δ	= angle of deviation.
δ	= rotational aberration coefficient.
δ	= aberration coefficient (coma).

<i>Symbol</i>	<i>Significance</i>
Δ	= finite increment of a variable.
${}^2\Delta, {}^3\Delta, \dots$, etc.	= second, third, etc., difference between values of a variable.
ϵ	= aberration coefficient (coma).
ζ	= aberration coefficient (spherical).
$\theta, \theta_1, \theta_2, \theta_3$	= angular displacement.
κ	= ray path parameter.
λ, λ_0	= wave-length.
μ	= coil factor.
μ	= refractive index.
ν	= frequency.
ξ	= ray path parameter.
π	= ratio of circumference to diameter of circle.
ρ	= charge density per unit volume.
ρ	= ray path parameter.
ρ, ρ_1	= radius of curvature.
ρ, ρ_s	= total radial aberration coefficient.
σ	= area.
σ	= component of ray (in Hamiltonian theory).
τ	= component of ray (in Hamiltonian theory).
u	= component of ray (in Hamiltonian theory).
ϕ, ϕ_D	= angle, or angular displacement.
ϕ	= potential function.
Φ	= magnetic scalar potential.
χ	= angle; angular displacement.
ψ	= potential function.
ψ	= angle.
ψ, ψ_n, ψ_{rz}	= magnetic flux.
	= aberration parameters.
∇	= $(\partial/\partial x + \partial/\partial y + \partial/\partial z)$ of a quantity $(S) = \text{grad } S$.
∇^2	= $(\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2)$ of a quantity = $\text{div grad } S$.
∞	= infinity.
\propto	= proportionality sign.
$ r $	= magnitude of r .

INDEX

(Main treatment of an item is indicated by heavy type: e.g. 111.)

- Abbe condition, 137, 200, 205.
Aberration: chromatic, 68, **138**, 205, 209, 211; of electron lenses, 12, 113, 250; of electron microscope, 205, 209, 211; of magnetic lenses, 100, 125, 259, 262; spherical, **121**, 205, 209, 211, 259.
Aperture lens, 42, **56**, 230; focal length of, 59; combinations of, 61.
Applegate diagram, 239-42.
Ardonne, v., 172, 203, 205, 207, 210, 211, 212, 214.
Astigmatism, **133**, 259.
Aston, 234, 235, 237.
Auto-electronic emission, 162.
Bachman, 210, 211, 214.
Bainbridge, 236, 237.
Baird, 166.
Beam power valve, 232.
Bedford, 76.
Benjamin, 163.
Bergmann, 91, 112.
Bertram, 20, 41.
Beta-ray spectrometry, 2, 87, **215**, 233.
Betatron, 2, **226**.
Birge, 264.
Boersch, 212, 214.
Borries, 147.
Bouwers, 80, 112.
Brainerd, 236, 248.
Braun tube, **3**, 159.
Brillouin, 220, 222, 236.
Britton, 236.
Brockway, 199.
de Broglie, 1, 2, 14, 194, 197, 249.
Brüche, 248, 263.
Bull, 236.
Burch, 191, 192.
Burton, 214.
Busch, 1, 15, 77, 86, 87, 91, 100, 107, 112, 263.
Calbick, 76, 230.
Calutron, 236.
Campbell, 172.
Cardinal points, **44**, 49, 53, 95-7.
Cathode-ray tube, 2, 3, 4, 6, 12, 68, 155, **173**; deflexion system, 174, **184**; double-beam, **175**; high-voltage, 14, **190**.
Chadwick, 236.
Christopherson, 41.
Chromatic aberration, 68, **138**, 205, 209, 211.
Cochrane, 214.
Cockcroft, 229, 236.
Coil factor, 89.
Compa, **136**, 259.
Conformal mapping, 20.
Cosslett, 199, 214, 218, 236.
Crookes, 3.
Cross-over, 70, 156, 158.
Crowther, 172.
Curvatura of field, **131**, 259.
Cyclotron, 2, 5, 12, 222.
Cylindrical lens, 230.
Davissou, 76, 194, 230.
Deflexion: electrostatic, 174, **184**; magnetic, 174, 179, **186**; system, 183, **184**.
Dempster, 235, 236, 237.
Deutsch, 218, 236.
Dieckmann, 167.
Diels, 121, 147, 263.
Diffraction, electron, 2, **194**.
Dirac, 150, 153.
Distortion, **130**, 259.
Dosse, 93, 94, 104, 112, 147, 163, 172.
Du Bridge, 172.
Ehrke, 162, 172.
Eiconal, 249, 253, 255, 258.
Electrolytic tank, **25**, 32.
Electron diffraction, 2, **194**; gun, **155**, 176, 179; microscopy, 2, 14, 18, 68, 90, 93, 163, **199**; mirrors, **71**, 125, 140; multipliers, 168, **169**.
Electron-volt, 13.
Elliott, 218, 236.
Ellis, 236.
e/m, determination of, 7, 9.
Emitron, 167, 177, **181**.
Epstein, 49, 55, 76, 263.
Equipotential surfaces, 9, 10, 11; in aperture lens, 57; in immersion lens, 69; in symmetrical lens, 62; in two-tube lens, 28.
Evans, 218, 236.
Farnsworth, 25, 80, 112, 136, 138, 158, 167, 171, 172, 177.
Fermat, 143, 249-52.
Fermi, 150, 153.
Field, 41, 55, 76.
Field distribution: in immersion lens, 69, 70; in magnetic lens, **87**, 93; in symmetrical lens, 62, 66; in two-tube lens, 28, 67.
Field emission, 152, **160**, 190, 192; microscope, 18, **163**, 212.
Finch, 192, 198, 214.
Fizeau effect, 105.

- Focal length: of combined lenses, 110;
of electrostatic lenses, 60, 65, 70; of
magnetic lenses, 82-4.
- Fortescue, 25.
- Fowler, 162, 172.
- Fringing, 6, **186**, 189.
- Funk, 250, 254, 262.
- Gabor, 32, 41, 105, 143, 147, 236, 262.
- Gas focusing, 159.
- Germer, 194, 197.
- de Gier, 193.
- Gill, 236.
- Glaser, 1, 79, 90, 91, 92, 94, 96, 97, 100,
107, 112, 115, 126-9, 141, 147, 249, 250,
252, 253-5, 258-62.
- Goddard, 35, 41, 98-100, 104, 112, 125.
- Gravitational model, 38 et seq.
- Gray, 128, 147, 250, 254, 262.
- Grösser, 69, 76.
- Haefl, 246, 248.
- Hahn, 247, 248.
- Hamilton, 1, 2, 249, 250, 253, 255, 262.
- Hamilton-Jacobi equation, 249, 250, 254.
- Harvey, 236, 248.
- Heidenreich, 203, 214.
- Hell, 167.
- Henley, 193.
- Henneberg, 147.
- Hillier, 206, 209, 213, 214.
- Holst, 166.
- Hottenroth, 76.
- Hughes, 172.
- Hull, 220, 236.
- Hutter, 112.
- Iams, 181, 193.
- Iconoscope, 167, **177**.
- Image-converter: electrostatic, 71, 136,
138, **166**; magnetic, 80, 136, 138, **167**,
177.
- Immersion lens: electrostatic, 43, **68**,
155, 173, 176; magnetic, **89**, 202, 208.
- Induction accelerator, 226.
- Induction, magnetic, 4.
- Jenkins, 163, 172.
- Johannson, 67, 69, 76, 87.
- Johnson, 163, 172.
- Jordan, 236, 237.
- Kapitza, 80.
- Kerst, 227, 228, 236.
- Kinescope, 175.
- Klanfer, 24, 41.
- Kleen, 236.
- Klemperer, 30, 37, 41, 49, 55, 76, 87, 93,
98-100, 104, 112, 123, 125, 135, 147,
214, 216, 236, 263.
- Kleynen, 39, 41.
- Klystron, 236, **242**; reflex, 246.
- Knecht, 87.
- Knoll, 89, 201, **206**, 214, 231, 236.
- Kohl, 214.
- Lagrange, 46, 137.
- Lammel, 112.
- Lamont, 248.
- Langmuir, 34, 41, 156, 172.
- Laplace, 16-18, 58, 105.
- Law, 159, 172.
- Lawrence, 223, 236.
- Lens: aperture, 42, **56**, 230; cardinal
points, 44, 48; cylindrical, 230; gauze,
16; ideal, 144; immersion, 43, **68**, **89**,
155, 173, 176; magnetic, **77**, 89; multi-
tube, 55; symmetrical, 42, **62**, 210-13;
two-tube, 16, 18, 28, 36-8, 42, **52**, 67,
155, 166; univoltage, 42, **62**, 210-13.
- Livingston, 223, 224, 236.
- Lubszynski, 181, 193.
- McGee, 181, 193.
- MacGregor-Morris, 193.
- Magnetic field measurement, 93.
- Magnetic lens: electrostatic equivalent,
104; immersion, **91**, 202, 208; short,
77, **80**, 158; solenoidal, **77**, 167;
spectrometer, 216.
- Magnetron, 5, 9, **219**, 236.
- Mahl, 210, 211, 214.
- Maloff, 55, 76, 263.
- Malter effect, 169, 172.
- Martin, 206, 214.
- Marton, 93, 112, 206, 214.
- Mass spectrographs, **233**.
- Megaw, 236.
- Meridional potential, 103, 249, 260, 262.
- Metcalf, 247, 248.
- Microscopes: electron, 2, 14, **199**; elec-
trostatic, 68, **210**; electrostatic pro-
jection, 212; magnetic, 90, **93**, **206**;
point-projector, 18, **163**, 212; resolv-
ing power, 199; scanning, 212.
- Mid-focal length, 47, 91, 94.
- Millor, 112, 193.
- Millman and Seely, 193, 263.
- Mirrors, electron, **71**, 125, 140.
- Morton, 70, 76, 112, 135, 147, 172, 176,
189, 193, 263.
- Motz, 22, 24, 41, 128.
- Müller, 163, 172.
- Multipliers, electron, 9, 12, **169**.
- Myers, 263.
- Nergaard, 248.
- Newton, 46.
- Nicoll, 72, 73, 76.
- Nodal points, **46**.
- Nordheim, 153, 162, 172.
- Oersted, 101.
- Okabe, 222.
- Orthicon, 167, 177, **181**.
- Oscillograph: cathode-ray, 12, 68, 155,
173; high-speed, **190**.
- Parr, 193.
- Pauli, 150.
- Peck, 203, 214.
- Perrin, 3.

- Petzval condition, 135.
 Phase-focusing, 13, 238.
 Photo-electric effect, 7, 9, 178, 181, 182.
 Photo-emission, 152, 164.
 Picht, 111, 249-63 *passim*.
 Pidduck, 236.
 Pierce, 156, 157, 172, 193.
 Planck, 14, 152, 164.
 Plass, 66, 67, 76, 128, 129, 147.
 Poisson, 17.
 Polotovskii, 62.
 Post-acceleration, 177.
 Prebus, 206, 214.
 Preston, 91, 112.
 Principal points, 45, 49, 53, 95-7.
 Radio valves, 229.
 Rajchmann, 39, 41, 171, 172.
 Ramberg, 70, 76, 112, 115, 129, 135, 142, 147, 211.
 Ramo, 210, 211, 214.
 Ray equation: combined fields, 106; electrostatic, 36, 64; magnetic, 84, 97; special solutions, 65, 109, 118.
 Ray-tracer, automatic, 32.
 Ray-tracing: electrostatic, 30; magnetic, 97.
 Rebsch, 127, 129, 142, 147.
 Recknagel, 76, 147, 248.
 Refractive index (equivalent), 10, 12.
 Relativistic mass, 13.
 Relaxation methods, 22.
 Resolving power, 199.
 Reyner, 193.
 Rhumbatron, 243.
 Richardson's equation, 153, 161, 165, 172.
 Ritchie, 172.
 Robinson, 193.
 Rodda, 181, 193, 236.
 Rose, 181, 193.
 Rotation of image, 85, 108, 181.
 Rothe, 236.
 Rubber model, 38.
 Ruska, 89, 92, 94, 95, 112, 125, 147, 156, 201, 206, 214.
 Rutherford, 236.
 Salinger, 32.
 Sae-Kulczycki, 35.
 Sass, 214.
 Schade, 236.
 Scherzer, 1, 2, 24, 65-7, 76, 111, 115, 126-8, 147, 250, 258, 259, 261, 263.
 Schneider, 129.
 Schottky effect, 161, 172.
 Schrödinger, 249.
 Secondary emission, 152, 168, 180.
 Seidel, 257, 258.
 Shockley, 163, 172.
 Shoenberg, 236.
 Siday, 97, 112.
 Siegman (K.), 91, 112, 218, 236.
 Slack, 162, 172.
 Slepian, 169.
 Smyth, 237.
 Snell's Law, 10, 29, 30, 104.
 Snyder, 213, 214.
 Solenoidal lens, 77, 167.
 Southwell, 22, 41.
 Space charge, 117, 142.
 Spangenberg, 41.
 Spherical aberration, 121, 205, 209, 211, 259.
 Starks, 214.
 Stoney, Johnstone, 3.
 Swinton, Campbell, 178, 193.
 Symmetrical lens, 42, 62, 66, 210-13; mirror, 74, 125, 140.
 Sygne, 250, 257, 262.
 Television, 70, 80, 156, 159, 166, 175, 177, 183.
 Tetrode, beam, 232.
 Thermionic emission, 151, 153; valves, 2, 18, 229.
 Thompson (H. C.), 236.
 Thomson (G. P.), 195, 197, 214.
 Thomson (J. J.), 3, 4, 7, 9, 14, 169, 233, 234, 235, 237.
 Townsend, 3.
 Trajectory plotting, 30, 34, 97.
 Tricker, 80, 112.
 Triode, 231.
 Tunnel effect, 162.
 Two-cylinder lens, 16, 18, 28, 36-8, 42, 52, 67, 155, 166; mirror, 72.
 Univoltage lens, 42, 62, 210-13.
 Vance, 209, 214.
 Varian, 243, 248.
 Vector potential, 97, 101.
 Velocity modulation, 238.
 Voit, 121, 135, 142, 147.
 Wallauschek, 91, 112.
 Walton, 227.
 Wavelength, electronic, 14, 194, 197, 201, 264-5.
 Webster, 248.
 Wehnelt, 4, 44, 69, 192, 198.
 Weiss, 171, 172.
 Wendt, 121, 147, 263.
 Whelpton, 191, 192.
 Wideröe, 227.
 Wiechert, 4, 77.
 Wierl, 199.
 Wilson (A. H.), 172.
 Wilson (W.), 193.
 Witcher, 80, 112.
 Wright, 30, 37, 41, 49, 76, 123, 135, 147.
 X-rays, 2, 7, 162, 164, 195, 196, 197, 201, 213.
 Young, 236, 237.
 Zworykin, 4, 39, 41, 44, 52, 70, 80, 136, 138, 155, 166, 169, 171, 172, 176, 178, 180, 189, 193, 206, 213, 214, 263.

